

SPRINGER
REFERENCE

George Z. Voyiadjis
Editor

Handbook of
Nonlocal
Continuum
Mechanics for
Materials and
Structures

 Springer

Handbook of Nonlocal Continuum Mechanics for Materials and Structures

George Z. Voyiadjis
Editor

Handbook of Nonlocal Continuum Mechanics for Materials and Structures

With 572 Figures and 59 Tables

 Springer

Editor

George Z. Voyiadjis
Department of Civil and Environmental Engineering
Louisiana State University
Baton Rouge, LA, USA

ISBN 978-3-319-58727-1 ISBN 978-3-319-58729-5 (eBook)
ISBN 978-3-319-58728-8 (print and electronic bundle)
<https://doi.org/10.1007/978-3-319-58729-5>

Library of Congress Control Number: 2018960886

© Springer Nature Switzerland AG 2019

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This handbook discusses the integral and gradient formulations of nonlocality, computational aspects, micromechanical considerations, and comparison of approaches and emphasizes recent developments in the bridging of material length and time scales. The contributions in this handbook are on nonlocal continuum plasticity in terms of the experimental, theoretical, and numerical investigations. This handbook presents a comprehensive treatment of the most important areas of nonlocality (integral and gradient) of time-dependent inelastic deformation behavior and heat transfer responses. The following aspects of the advanced material modeling are presented: enhanced (generalized) continuum mechanics, microscopic mechanisms and micro-mechanical aspects responsible for size effect and micro-scale heat transfer, thermodynamic framework, and multiscale computational aspects with detailed nonlocal computational algorithms in the context of finite element. Measures for length scales are introduced together with their appropriate evolution relations.

The work addresses the thermal and mechanical responses of small-scale metallic compounds under fast transient processes based on small and large deformation framework. This handbook presents a comprehensive treatment of the most important areas of nonlocality (integral and gradient) of time-dependent inelastic deformation behavior and heat transfer responses.

From the experimental aspects, a wide spectrum of materials are included in these chapters. For the case of glassy polymers, their nanostructural responses and nanoindentation measurements were discussed in detail by Voyiadjis et al. For metals, the hydrogen embrittlement cracking was discussed by Yonezu and Chen, whereas its size effect of nanoindentation was elucidated by Voyiadjis et al. Composites and their nonlocal behaviors have received particular attention, including the studies of their cracking initiation (Xu et al.), interface stability (Meng et al.), buckling (Chen et al.), and dynamic properties (Tomar et al.). Indentation and its nonlocal characteristics were also emphasized through modeling (Mills et al., Liu et al.) and fatigue testing (Xu et al.).

The micromorphic approach has aroused strong interest from the materials science and computational mechanics communities because of its regularization power in the context of softening plasticity and damage. The micromorphic and Cosserat theories in gradient plasticity are introduced and analyzed in these chapters to address the instabilities in the materials and structures. The micromorphic

approach for the gradient continuum plasticity/damage and crystal plasticity is elucidated by Forest et al. The micropolar theory for the crystal plasticity is also introduced (Mayeur et al.). The micromorphic and Cosserat approach are applied to localization in geomaterials (Stefanou et al.) and dispersion of waves in metamaterials (Madeo et al.) in these chapters.

The section on “Mathematical Methods in Nonlocal Continuum Mechanics” combines the original concepts for the description of nonlocal effects, including applications to nonstandard mathematical models. Both quasi-static and dynamic processes are included, together with micro- and macro-level of description, and furthermore multiphysics (in the sense of, e.g., thermomechanical interaction) is covered accordingly. In the chapter by Lazopoulos et al., the fractional calculus is applied to obtain the space-fractional nonlocal continuum mechanics formulation and then applied to the analysis of fractional Zener viscoelastic model. The contribution by Ostoja-Starzewski et al. includes modeling of fractal materials utilizing fractional integrals and application of homogenized continuum mechanics together with the framework of calculus in non-integer dimensional spaces. The work by Sumelka et al. presents implicit time nonlocal modeling of metallic materials including damage anisotropy, and furthermore stress-fractional extensions are suggested. Next, the chapter by Tarasov includes modeling of physical lattices with long-range interactions utilizing exact fractional-order difference operators. Finally, in the chapter by Voyiadjis et al., the strain gradient plasticity is presented with the appropriate flow rules of the grain interior and grain boundary areas within the thermodynamically consistent framework and applied for modeling metallic materials. One can recapitulate that the overall section contents open new fields of investigations of mathematical models for bodies exhibiting strong scale effect.

The section on “Computational Modeling for Gradient Plasticity in Both Temporal and Spatial Scales” introduces a variety of numerical examples for the gradient-enhanced plasticity. The failure mechanisms of metallic materials for high-velocity impact loading are simulated based on the nonlocal approach (Voyiadjis et al.). The gradient plasticity is combined to micro-/mesoscale crystal plasticity (Yalcinkaya et al., Ozdemir et al.) and fracture mechanics (Lancioni et al.). The transverse vibration of microbeams and axial vibration of micro-rods are studied by Civalek et al. using the strain gradient plasticity theory. In addition, in this section, the temperature-driven ductile-to-brittle transition fracture in Ferritic steels is modeled and analyzed (Deliktaş et al.).

Nonlocal peridynamic models for damage and fracture are introduced and analyzed in these chapters. The theory guaranteeing the existence of solution for continuum models of fracture evolution is essential for the development of mesh-independent discretizations. A bond is exhibited based on peridynamic models with damage and softening and shows that they are well-posed evolutions both over the space of Holder continuous functions and Sobolev functions (Jha and Lipton). This feature is used to develop numerical convergence rates in space and time for the associated finite element and finite-difference schemes (Jha and Lipton). These are the first convergence rates for numerical schemes applied to fully nonlinear and nonlocal peridynamic models. A more general state-based peridynamic model is

developed for free fracture evolution. In the limit of vanishing nonlocal interaction, this model is shown to converge to the equation of elastic momentum balance away from the crack set (Said et al.). In the final chapter, a state-based and history-dependent dynamic damage model is developed. Several numerical examples are provided illustrating the theory (Said et al.).

This handbook integrates knowledge from the theoretical, numerical, and experimental areas of nonlocal continuum plasticity. This book is focused mainly for graduate students of nonlocal continuum plasticity, researchers in academia and industry who are active or intend to become active in this field, and practicing engineers and scientists who work in this topic and would like to solve problems utilizing the tools offered by nonlocal mechanics. This handbook should serve as an excellent text for a series of graduate courses in mechanical engineering, civil engineering, materials science, engineering mechanics, aerospace engineering, applied mathematics, applied physics, or applied chemistry.

This handbook is basically intended as a textbook for university courses as well as a reference for researchers in this field. It will serve as a timely addition to the literature on nonlocal mechanics and will serve as an invaluable resource to members of the international scientific and industrial communities.

It is hoped that the reader will find this handbook a useful resource as he/she progresses in their study and research in nonlocal mechanics. Each of the individual sections of this handbook could be considered as a compact self-contained mini-book right under its own title. However, these topics are presented in relation to the basic principles of nonlocal mechanics.

What is finally presented in the handbook is the work contributed by celebrated international experts for their best knowledge and practices on specific and related topics in nonlocal mechanics.

The editor would like to thank all the contributors who wrote chapters for this handbook. Finally, the editor would like to acknowledge the help and support of his family members and the editors at Springer who made this handbook possible.

Baton Rouge, USA
December 2018

Dr. George Z. Voyiadjis

Contents

Volume 1

Part I Nanoindentation for Length Scales	1
1 Size Effects and Material Length Scales in Nanoindentation for Metals	3
George Z. Voyiadjis and Cheng Zhang	
2 Size Effects During Nanoindentation: Molecular Dynamics Simulation	39
George Z. Voyiadjis and Mohammadreza Yaghoobi	
3 Molecular Dynamics-Decorated Finite Element Method (MDeFEM): Application to the Gating Mechanism of Mechanosensitive Channels	77
Liangliang Zhu, Qiang Cui, Yilun Liu, Yuan Yan, Hang Xiao, and Xi Chen	
4 Spherical Indentation on a Prestressed Elastic Coating/Substrate System	129
James A. Mills and Xi Chen	
5 Experimentation and Modeling of Mechanical Integrity and Instability at Metal/Ceramic Interfaces	153
Wen Jin Meng and Shuai Shao	
6 Uniqueness of Elastoplastic Properties Measured by Instrumented Indentation	211
L. Liu, Xi Chen, N. Ogasawara, and N. Chiba	
7 Helical Buckling Behaviors of the Nanowire/Substrate System	241
Youlong Chen, Yilun Liu, and Xi Chen	
8 Hydrogen Embrittlement Cracking Produced by Indentation Test	289
Akio Yonezu and Xi Chen	

9	Continuous Stiffness Measurement Nanoindentation Experiments on Polymeric Glasses: Strain Rate Alteration	315
	George Z. Voyiadjis, Leila Malekmotiei, and Aref Samadi-Dooki	
10	Shear Transformation Zones in Amorphous Polymers: Geometrical and Micromechanical Properties	333
	George Z. Voyiadjis, Leila Malekmotiei, and Aref Samadi-Dooki	
11	Properties of Material Interfaces: Dynamic Local Versus Nonlocal	361
	Devendra Verma, Chandra Prakash, and Vikas Tomar	
12	Nanostructural Response to Plastic Deformation in Glassy Polymers	377
	George Z. Voyiadjis and Aref Samadi-Dooki	
13	Indentation Fatigue Mechanics	401
	Baoxing Xu, Xi Chen, and Zhufeng Yue	
14	Crack Initiation and Propagation in Laminated Composite Materials	433
	Jun Xu and Yanting Zheng	
Part II	Micromorphic and Cosserat in Gradient Plasticity for Instabilities in Materials and Structures	497
15	Micromorphic Approach to Gradient Plasticity and Damage	499
	Samuel Forest	
16	Higher Order Thermo-mechanical Gradient Plasticity Model: Nonproportional Loading with Energetic and Dissipative Components	547
	George Z. Voyiadjis and Yooseob Song	
17	Micropolar Crystal Plasticity	595
	J. R. Mayeur, D. L. McDowell, and Samuel Forest	
18	Micromorphic Crystal Plasticity	643
	Samuel Forest, J. R. Mayeur, and D. L. McDowell	
19	Cosserat Approach to Localization in Geomaterials	687
	Ioannis Stefanou, Jean Sulem, and Hadrien Rattiez	
20	Dispersion of Waves in Micromorphic Media and Metamaterials	713
	Angela Madeo and Patrizio Neff	

Volume 2

Part III	Mathematical Methods in Nonlocal Continuum Mechanics	741
21	Implicit Nonlocality in the Framework of the Viscoplasticity	743
	Wojciech Sumelka and Tomasz Łodygowski	
22	Finite Element Analysis of Thermodynamically Consistent Strain Gradient Plasticity Theory and Applications	781
	George Z. Voyiadjis and Yooseob Song	
23	Fractional Nonlocal Continuum Mechanics and Microstructural Models	839
	Vasily E. Tarasov	
24	Fractional Differential Calculus and Continuum Mechanics	851
	K. A. Lazopoulos and A. K. Lazopoulos	
25	Continuum Homogenization of Fractal Media	905
	Martin Ostoja-Starzewski, Jun Li, and Paul N. Demmie	
Part IV	Computational Modeling for Gradient Plasticity in Both Temporal and Spatial Scales	937
26	Modeling High-Speed Impact Failure of Metallic Materials: Nonlocal Approaches	939
	George Z. Voyiadjis and Babür Deliktaş	
27	Strain Gradient Plasticity: Deformation Patterning, Localization, and Fracture	971
	Giovanni Lancioni and Tuncay Yalçinkaya	
28	Strain Gradient Crystal Plasticity: Thermodynamics and Implementation	1001
	Tuncay Yalçinkaya	
29	Strain Gradient Crystal Plasticity: Intergranular Microstructure Formation	1035
	İzzet Özdemir and Tuncay Yalçinkaya	
30	Microplane Models for Elasticity and Inelasticity of Engineering Materials	1065
	Ferhun C. Caner, Valentín de Carlos Blasco, and Mercè Ginjaume Egido	

31	Modeling Temperature-Driven Ductile-to-Brittle Transition Fracture in Ferritic Steels	1099
	Babür Deliktaş, Ismail Cem Turtuk, and George Z. Voyiadjis	
32	Size-Dependent Transverse Vibration of Microbeams	1123
	Ömer Civalek and Bekir Akgöz	
33	Axial Vibration of Strain Gradient Micro-rods	1141
	Ömer Civalek, Bekir Akgöz, and Babür Deliktaş	
Part V	Peridynamics	1157
34	Peridynamics: Introduction	1159
	S. A. Silling	
35	Recent Progress in Mathematical and Computational Aspects of Peridynamics	1197
	Marta D’Elia, Qiang Du, and Max Gunzburger	
36	Optimization-Based Coupling of Local and Nonlocal Models: Applications to Peridynamics	1223
	Marta D’Elia, Pavel Bochev, David Littlewood, and Mauro Perego	
37	Bridging Local and Nonlocal Models: Convergence and Regularity	1243
	Mikil D. Foss and Petronela Radu	
38	Dynamic Brittle Fracture from Nonlocal Double-Well Potentials: A State-Based Model	1265
	Robert Lipton, Eyad Said, and Prashant K. Jha	
39	Nonlocal Operators with Local Boundary Conditions: An Overview	1293
	Burak Aksoylu, Fatih Celiker, and Orsan Kilicer	
40	Peridynamics and Nonlocal Diffusion Models: Fast Numerical Methods	1331
	Hong Wang	
41	Peridynamic Functionally Graded and Porous Materials: Modeling Fracture and Damage	1353
	Ziguang Chen, Sina Niazi, Guanfeng Zhang, and Florin Bobaru	
42	Numerical Tools for Improved Convergence of Meshfree Peridynamic Discretizations	1389
	Pablo Seleson and David J. Littlewood	
43	Well-Posed Nonlinear Nonlocal Fracture Models Associated with Double-Well Potentials	1417
	Prashant K. Jha and Robert Lipton	

44	Finite Differences and Finite Elements in Nonlocal Fracture Modeling: A Priori Convergence Rates	1457
	Prashant K. Jha and Robert Lipton	
45	Dynamic Damage Propagation with Memory: A State-Based Model	1495
	Robert Lipton, Eyad Said, and Prashant K. Jha	
Index	1525

About the Editor



George Z. Voyiadjis is the Boyd Professor at the Louisiana State University in the Department of Civil and Environmental Engineering. This is the highest professorial rank awarded by the Louisiana State University System. He is also the holder of the Freeport-McMoRan Endowed Chair in Engineering. He joined the faculty of Louisiana State University in 1980. He is currently the Chair of the Department of Civil and Environmental Engineering. He holds this position since February of 2001. He currently also serves since 2012 as the Director of the Louisiana State University Center for GeoInformatics (LSU C4G; <http://c4gnet.lsu.edu/c4g/>).

Voyiadjis is a Foreign Member of both the Polish Academy of Sciences, Division IV (Technical Sciences), and the National Academy of Engineering of Korea. He is the recipient of the 2008 Nathan M. Newmark Medal of the American Society of Civil Engineers and the 2012 Khan International Medal for outstanding lifelong contribution to the field of plasticity. He was also the recipient of the Medal for his significant contribution to Continuum Damage Mechanics, presented to him during the Second International Conference on Damage Mechanics (ICDM2), Troyes, France, July, 2015. This is sponsored by the *International Journal of Damage Mechanics* and is held every 3 years.

Voyiadjis was honored in April of 2012 by the International Symposium on “Modeling Material Behavior at Multiple Scales” sponsored by Hanyang University, Seoul, Korea, chaired by T. Park and X. Chen (with a dedicated special issue in the *Journal of Engineering Materials and Technology* of the ASME). He was also honored by an International Mini-Symposium on “Multiscale and Mechanism Oriented Models:

Computations and Experiments” sponsored by the International Symposium on Plasticity and Its Current Applications, chaired by V. Tomar and X. Chen, in January 2013.

He is a Distinguished Member of the American Society of Civil Engineers; Fellow of the American Society of Mechanical Engineers, the Society of Engineering Science, the American Academy of Mechanics and the Engineering Mechanics Institute of ASCE; and Associate Fellow of the American Institute of Aeronautics and Astronautics. He was on the Board of Governors of the Engineering Mechanics Institute of the American Society of Civil Engineers, and Past President of the Board of Directors of the Society of Engineering Science. He was also the Chair of the Executive Committee of the Materials Division (MD) of the American Society of Mechanical Engineers. Dr. Voyiadjis is the Founding Chief Editor of the Journal of Nanomechanics and Micromechanics of the ASCE and is on the editorial board of numerous engineering journals. He was also selected by Korea Science and Engineering Foundation (KOSEF) as one of the only two World-Class University foreign scholars in the area of civil and architectural engineering to work on nanofusion in civil engineering. This is a multimillion research grant.

Voyiadjis' primary research interest is in plasticity and damage mechanics of metals, metal matrix composites, polymers, and ceramics with emphasis on the theoretical modeling, numerical simulation of material behavior, and experimental correlation. Research activities of particular interest encompass macro-mechanical and micro-mechanical constitutive modeling, experimental procedures for quantification of crack densities, inelastic behavior, thermal effects, interfaces, damage, failure, fracture, impact, and numerical modeling.

Dr. Voyiadjis' research has been performed on developing numerical models that aim at simulating the damage and dynamic failure response of advanced engineering materials and structures under high-speed impact loading conditions. This work will guide the development of design criteria and fabrication processes of high-performance materials and structures under severe loading conditions. Emphasis is placed

on survivability area that aims to develop and field a contingency armor that is thin and lightweight, but with a very high level of an overpressure protection system that provides low penetration depths. The formation of cracks and voids in the adiabatic shear bands, which are the precursors to fracture, is mainly investigated.

He has 2 patents, over 320 refereed journal articles, and 19 books (11 as editor) to his credit. He gave over 400 presentations as plenary, keynote, and invited speaker as well as other talks. Over 62 graduate students (36 Ph.D.) completed their degrees under his direction. He has also supervised numerous postdoctoral associates. Voyiadjis has been extremely successful in securing more than \$25.0 million in research funds as a principal investigator/investigator from the National Science Foundation, the Department of Defense, the Air Force Office of Scientific Research, the Department of Transportation, and major companies such as IBM and Martin Marietta.

He has been invited to give plenary presentations and keynote lectures in many countries around the world. He has also been invited as guest editor in numerous volumes of the *Journal of Computer Methods in Applied Mechanics and Engineering*, *International Journal of Plasticity*, *Journal of Engineering Mechanics of the ASCE*, and *Journal of Mechanics of Materials*. These special issues focus in the areas of damage mechanics, structures, fracture mechanics, localization, and bridging of length scales.

He has extensive international collaborations with universities in France, the Republic of Korea, and Poland.

Associate Editors



Xi Chen

Department of Earth and Environmental Engineering
Columbia Nanomechanics Research Center
Columbia University
New York, NY, USA



Samuel Forest

Centre des Materiaux
Mines ParisTech CNRS
PSL Research University
CNRS UMR 7633
Evry Cedex, Paris, France



Wojciech Sumelka

Institute of Structural Engineering
Poznan University of Technology
Poznan, Poland



Babür Deliktaş
Faculty of Engineering-Architecture
Department of Civil Engineering
Uludag University
Görükle, Bursa, Turkey



Michael L. Parks
Center for Computing Research
Sandia National Laboratories
Albuquerque, NM, USA

Contributors

Bekir Akgöz Civil Engineering Department, Division of Mechanics, Akdeniz University, Antalya, Turkey

Burak Aksoylu Department of Mathematics, Wayne State University, Detroit, MI, USA

Florin Bobaru Mechanical and Materials Engineering, University of Nebraska–Lincoln, Lincoln, NE, USA

Pavel Bochev Center for Computing Research, Sandia National Laboratories, Albuquerque, NM, USA

Ferhun C. Caner School of Industrial Engineering, Institute of Energy Technologies, Universitat Politècnica de Catalunya, Barcelona, Spain

Department of Materials Science and Metallurgical Engineering, Universitat Politècnica de Catalunya, Barcelona, Spain

Fatih Celiker Department of Mathematics, Wayne State University, Detroit, MI, USA

Xi Chen Department of Earth and Environmental Engineering, Columbia Nanomechanics Research Center, Columbia University, New York, NY, USA

Youlong Chen International Center for Applied Mechanics, State Key Laboratory for Strength and Vibration of Mechanical Structures, School of Aerospace, Xi'an Jiaotong University, Xi'an, China

Ziguang Chen Department of Mechanics, Huazhong University of Science and Technology, Wuhan, Hubei Sheng, China

Hubei Key Laboratory of Engineering, Structural Analysis and Safety Assessment, Wuhan, China

N. Chiba National Defense Academy of Japan, Yokosuka, Japan

Ömer Civalek Civil Engineering Department, Division of Mechanics, Akdeniz University, Antalya, Turkey

Qiang Cui Department of Chemistry and Theoretical Chemistry Institute, University of Wisconsin-Madison, Madison, WI, USA

Marta D'Elia Optimization and Uncertainty Quantification Department Center for Computing Research, Sandia National Laboratories, Albuquerque, NM, USA

Valentín de Carlos Blasco School of Industrial Engineering, Institute of Energy Technologies, Universitat Politècnica de Catalunya, Barcelona, Spain

Babür Deliktaş Faculty of Engineering-Architecture, Department of Civil Engineering, Uludag University, Bursa, Turkey

Paul N. Demmie Sandia National Laboratories, Albuquerque, NM, USA

Qiang Du Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY, USA

Samuel Forest Centre des Matériaux, Mines ParisTech CNRS, PSL Research University, CNRS UMR 7633, Evry Cedex, Paris, France

Mikil D. Foss Department of Mathematics, University of Nebraska-Lincoln, Lincoln, NE, USA

Mercè Ginjaume Egido School of Industrial Engineering, Institute of Energy Technologies, Universitat Politècnica de Catalunya, Barcelona, Spain

Max Gunzburger Department of Scientific Computing, Florida State University, Tallahassee, FL, USA

Prashant K. Jha Department of Mathematics, Louisiana State University, Baton Rouge, LA, USA

Orsan Kilicer Department of Mathematics, Wayne State University, Detroit, MI, USA

Giovanni Lancioni Dipartimento di Ingegneria Civile, Edile e Architettura, Università Politecnica delle Marche, Ancona, Italy

A. K. Lazopoulos Mathematical Sciences Department, Hellenic Army Academy, Vari, Greece

K. A. Lazopoulos National Technical University of Athens, Rafina, Greece

Jun Li Department of Mechanical Engineering, University of Massachusetts, Dartmouth, MA, USA

Robert Lipton Department of Mathematics and Center for Computation and Technology, Louisiana State University, Baton Rouge, LA, USA

David J. Littlewood Center for Computing Research, Sandia National Laboratories, Albuquerque, NM, USA

L. Liu Department of Mechanical and Aerospace Engineering, Utah State University, Logan, UT, USA

Yilun Liu International Center for Applied Mechanics, State Key Laboratory for Strength and Vibration of Mechanical Structures, School of Aerospace, Xi'an Jiaotong University, Xi'an, China

Tomasz Łodygowski Institute of Structural Engineering, Poznan University of Technology, Poznan, Poland

Angela Madeo SMS-ID, INSA-Lyon, Université de Lyon, Villeurbanne cedex, Lyon, France

Institut universitaire de France, Paris Cedex 05, Paris, France

Leila Malekmoiei Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge, LA, USA

J. R. Mayeur Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM, USA

D. L. McDowell Woodruff School of Mechanical Engineering, School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, GA, USA

Wen Jin Meng Department of Mechanical and Industrial Engineering, Louisiana State University, Baton Rouge, LA, USA

James A. Mills Department of Civil Engineering and Engineering Mechanics, Columbia University, New York, NY, USA

Patrizio Neff Fakultät für Mathematik, Universität Duisburg-Essen, Essen, Germany

Sina Niazi Mechanical and Materials Engineering, University of Nebraska–Lincoln, Lincoln, NE, USA

N. Ogasawara Department of Mechanical Engineering, National Defense Academy of Japan, Yokosuka, Japan

Martin Ostoja-Starzewski Department of Mechanical Science and Engineering, Institute for Condensed Matter Theory and Beckman Institute, University of Illinois at Urbana–Champaign, Urbana, IL, USA

İzzet Özdemir Department of Civil Engineering, İzmir Institute of Technology, İzmir, Turkey

Mauro Perego Center for Computing Research, Sandia National Laboratories, Albuquerque, NM, USA

Chandra Prakash School of Aeronautics and Astronautics, Purdue University, West Lafayette, IN, USA

Petronela Radu Department of Mathematics, University of Nebraska-Lincoln, Lincoln, NE, USA

Hadrien Rattiez Navier (CERMES), UMR 8205, Ecole des Ponts, IFSTTAR, CNRS, Champs-sur-Marne, France

Eyad Said Department of Mathematics, Louisiana State University, Baton Rouge, LA, USA

Aref Samadi-Dooki Computational Solid Mechanics Laboratory, Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge, LA, USA

Pablo Seleson Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA

Shuai Shao Department of Mechanical and Industrial Engineering, Louisiana State University, Baton Rouge, LA, USA

S. A. Silling Sandia National Laboratories, Albuquerque, NM, USA

Yooseob Song Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge, LA, USA

Ioannis Stefanou Navier (CERMES), UMR 8205, Ecole des Ponts, IFSTTAR, CNRS, Champs-sur-Marne, France

Jean Sulem Navier (CERMES), UMR 8205, Ecole des Ponts, IFSTTAR, CNRS, Champs-sur-Marne, France

Wojciech Sumelka Institute of Structural Engineering, Poznan University of Technology, Poznan, Poland

Vasily E. Tarasov Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, Moscow, Russia

Vikas Tomar School of Aeronautics and Astronautics, Purdue University, West Lafayette, IN, USA

Ismail Cem Turtuk Mechanical Design Department, Meteksan Defence, Ankara, Turkey

Department of Civil Engineering, Uludag Univeristy, Bursa, Turkey

Devendra Verma School of Aeronautics and Astronautics, Purdue University, West Lafayette, IN, USA

George Z. Voyiadjis Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge, LA, USA

Hong Wang Department of Mathematics, University of South Carolina, Columbia, SC, USA

Hang Xiao School of Chemical Engineering, Northwest University, Xi'an, China

Baoxing Xu Department of Mechanical and Aerospace Engineering, University of Virginia, Charlottesville, VA, USA

Jun Xu Department of Automotive Engineering, School of Transportation Science and Engineering, Beihang University, Beijing, China

Advanced Vehicle Research Center (AVRC), Beihang University, Beijing, China

Mohammadreza Yaghoobi Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge, LA, USA

Tuncay Yalçinkaya Aerospace Engineering Program, Middle East Technical University Northern Cyprus Campus, Guzelyurt, Mersin, Turkey

Department of Aerospace Engineering, Middle East Technical University, Ankara, Turkey

Yuan Yan School of Chemical Engineering, Northwest University, Xi'an, China

Akio Yonezu Department of Precision Mechanics, Chuo University, Tokyo, Japan

Zhufeng Yue Department of Engineering Mechanics, Northwestern Polytechnical University, Xi'an, Shaanxi, China

Cheng Zhang Medtronic, Inc., Tempe, AZ, USA

Guanfeng Zhang Mechanical and Materials Engineering, University of Nebraska–Lincoln, Lincoln, NE, USA

Yanting Zheng China Automotive Technology and Research Center, Tianjin, China

Liangliang Zhu Columbia Nanomechanics Research Center, Department of Earth and Environmental Engineering, Columbia University, New York, NY, USA

International Center for Applied Mechanics, State Key Laboratory for Strength and Vibration of Mechanical Structures, School of Aerospace, Xi'an Jiaotong University, Xi'an, China

Part I
Nanoindentation for Length Scales



Size Effects and Material Length Scales in Nanoindentation for Metals

1

George Z. Voyiadjis and Cheng Zhang

Contents

Introduction	4
Nonlocal Theory	6
Physically Based Material Length Scale	8
Determination of the Length Scales	11
Applications on Single Crystal, Polycrystalline, and Bicrystalline Metals	16
Sample Preparations	16
Temperature and Strain Rate Dependency on Single Crystal and Polycrystalline Metals	18
Influence of the Grain Boundaries on Bicrystal Metals	24
Summary and Conclusion	35
References	37

Abstract

In *nanoindentation* experiments at submicron indentation depths, the hardness decreases with the increasing indentation depth. This phenomenon is termed as the indentation size effect. In order to predict the indentation size effect, the classical continuum needs to be enhanced with the strain gradient plasticity theory. The strain gradient plasticity theory provides a nonlocal term in addition to the classical theory. A material length scale parameter is required to be incorporated into the constitutive expression in order to characterize the size effects in different materials. By comparing the model of hardness as a function

G. Z. Voyiadjis (✉)

Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge, LA, USA

e-mail: voyiadjis@eng.lsu.edu

C. Zhang

Medtronic, Inc., Tempe, AZ, USA

e-mail: cheng.zhang2@medtronic.com

© Springer Nature Switzerland AG 2019

G. Z. Voyiadjis (ed.), *Handbook of Nonlocal Continuum Mechanics for Materials and Structures*, https://doi.org/10.1007/978-3-319-58729-5_27

3

of the indentation depth with the nanoindentation experimental results, the length scale can be determined. Recent nanoindentation experiments on polycrystalline metals have shown an additional hardening segment in the hardness curves instead of the solely decreasing hardness as a function of the indentation depth. It is believed that the accumulation of dislocations near the grain boundaries during nanoindentation causes the additional increase in hardness. In order to isolate the influence of the grain boundary, bicrystal metals are tested near the grain boundary at different distances. The results show that the hardness increases with the decreasing distance between the indenter and the grain boundary, providing a new type of size effect. The length scales at different distances are determined using the modified model of hardness and the nanoindentation experimental results on bicrystal metals.

Keywords

Nanoindentation · Indentation size effect · Nonlocal theory · Strain gradient plasticity · Grain boundary · Bicrystal · Length scale

Introduction

It has been found in experiments at the microscopic scale that the mechanical response increases as the size of the specimen decreases. The phenomenon is termed as the size effect. Size effects have been reported in tests of metals in small scales such as microbending, microtorsion, and bulge tests (Fleck and Hutchinson 1997; Stolken and Evans 1998; Chen et al. 2007; Xiang and Vlassak 2005). It is believed that the size effects are attributed to nonuniform deformation in the small scale during the tests (Nye 1953). Geometrically necessary dislocations (GNDs) are generated in order to accommodate the deformation. The GNDs act as barriers of the generation of the statistically stored dislocations (SSDs), resulting in the increase of the mechanical response (De Guzman et al. 1993; Stelmashenko et al. 1993; Fleck et al. 1994; Ma and Clarke 1995). The classical continuum theory is only able to predict the macroscopic and therefore not capable in capturing the changes of mechanical responses. A nonlocal term is required in addition to the classical theory. In the study of the plastic deformation, the strain gradient plasticity theory is applied by adding a strain gradient term into the classical expression (Aifantis 1992; Zbib and Aifantis 1998). In order to characterize the influence of the strain gradient, a length scale parameter is required to be incorporated into the expression of the strain gradient plasticity theory. The length scale is an intrinsic parameter and each material has its unique length scale. Therefore, it becomes of great importance to determine the material intrinsic length scale. With the length scale determined, the strain gradient plasticity theory is able to predict the mechanical behaviors in both microscopic and macroscopic scales, bridging the gap between large and small scales.

When using the strain gradient plasticity theory to derive the expression of the length scales, material parameters are needed in order to represent the behavior for

different materials. In order to determine the length scale for a specific material, the materials need to be determined from experimental results. Nanoindentation experiments are believed to be the most effective technique to determine the material length scales (Begley and Hutchinson 1998). In nanoindentation at small depth, the hardness increases with the decreasing indentation depth (McElhane et al. 1998). The size effect encountered in nanoindentation is referred to the indentation size effect. With the use of the conical or Berkovich indenter, the indentation is said to be self-similar. The GND density can be calculated for a self-similar indent from the geometry of the indenter (Nix and Gao 1998). The hardness can be mapped from the stress-strain relation using Tabor's factor in macroscopic scale (Tabor 1951). In microscopic scales, the hardness is related to the dislocation density according to Taylor's hardening law. The hardness expression needs to incorporate the expression of the length scale in order to predict the indentation size effect. By comparing the experiments and the expression of hardness, the material parameters in the expression of the length scale can be determined.

More recent nanoindentation experiments in polycrystalline materials have shown that instead of the solely decreasing hardness with the increasing indentation depth, there is a hardening-softening phenomenon observed (Yang and Vehoff 2007; Voyiadjis and Peters 2010). The additional increase in hardness is believed to be due to the interaction between the dislocations generated by the penetration of the indenter and the grain boundaries. As there is difficulty for dislocations to transfer across the grain boundary, the dislocations accumulate near the grain boundaries. The local dislocation density increases due to the accumulation of dislocations, resulting in the additional increase in hardness. Once the dislocations start to move across the grain boundary at certain point, the dislocation density starts to decrease, which explains the softening effect following the hardening. In order to characterize the grain boundary effect as well the temperature and strain rate dependency of the length scale, a model considering the temperature and strain rate dependency of indentation size effect (TRISE) is developed (Voyiadjis et al. 2011). The length scale is written as a function of the temperature, plastic strain rate, and the grain size. By comparing the TRISE model with nanoindentation experiments at different temperatures, plastic strain rates, and grain sizes, the length scales can be determined.

In order to confirm the contribution of grain boundaries to the hardening-softening phenomenon, the investigation of the grain boundary is isolated through bicrystalline materials, where there is only one grain boundary. Nanoindentation experiments are conducted near the grain boundary at different distances between the indenter and the grain boundary (Voyiadjis and Zhang 2015; Zhang and Voyiadjis 2016). Single crystal behavior is observed for indents made with large distances to the grain boundary, without showing the additional hardening. Only the indents in close proximity of the grain boundary show the hardening-softening effect and there is a stronger hardening effect when the distance is smaller, providing a new type of size effect. The influence of the grain boundary on the hardness during nanoindentation is thus confirmed by experiments. The TRISE model is rewritten by replacing the grain size with the distance between the grain boundary

and the indenter. The length scales of bicrystalline materials can be determined through the comparisons between the developed TRISE model and the nanoindentation experiments in order to characterize the size effect regarding the grain boundary.

In this chapter, the dependencies of temperature, plastic strain rate, grain size, and the distance between the indents and the grain boundary are addressed during nanoindentation experiments on single crystal, polycrystalline, and bicrystalline materials. The TRISE model is developed and applied in order to predict the hardness as a function of the indentation depth. The equivalent plastic strain as a function of the indentation depth during nanoindentation is determined through the finite element method. ABAQUS/Explicit software is used in order to simulate the indentation problem. The materials tested are body-centered cubic (BCC) and face-centered cubic (FCC) materials. In order to show the distinct behaviors between FCC and BCC materials during the simulation, user material subroutines VUMAT are incorporated for BCC and FCC metals, respectively. The length scales are determined through the comparison between the nanoindentation experiments and the developed TRISE models.

Nonlocal Theory

According to the nonlocal theory, the material properties at a given material point are not only dependent on their local counterparts but also depend on the state of the neighboring space. While the classical continuum theory only provides the predictions of the local point, the nonlocal theory enhances the classical theory by giving a nonlocal gradient term. The nonlocal term is taken as the characterization of the interactions from the neighboring space. The nonlocal expression was incorporated through an integral form for the elastic models of materials (Kroner 1967; Eringen and Edelen 1972). In the integral format, the nonlocal measure \bar{A} at a given material point x is expressed through the weighted average of its local counterpart A over the surrounding volume V within a small distance d from the point x as follows:

$$\bar{A} = \frac{1}{V} \int_V w(d) A(x + d) dV \quad (1)$$

where $w(d)$ is a weight function that decays gradually with the distance d . It should be noted that there is a limit of the distance d , which is the internal characteristic length that shows the range of the influence of the nonlocal term.

The integration can be solved analytically for the elasticity problems. However, it is not practical to solve the integration for more complicated plasticity problems. Therefore, the integral expression needs to be simplified. The local counterpart A in the integral can be approximated using the Taylor's series expansion at $d = 0$ as follows (Muhlhaus and Aifantis 1991; Vardoulakis and Aifantis 1991):

$$A(x+d) = A(x) + \nabla A(x)d + \frac{1}{2!}\nabla^2 A(x)d^2 + \frac{1}{3!}\nabla^3 A(x)d^3 + \dots \quad (2)$$

where ∇^i denotes the i th order of the gradient operator. The expression can be further simplified considering only the isotropic behaviors. Therefore, the odd terms vanish in Eq. 2 and the nonlocal expression can be rewritten as follows:

$$\bar{A} = \frac{1}{V} \int_V w(d)A(x)dV + \frac{1}{2!V} \int_V w(d)\nabla^2 A(x)d^2dV \quad (3)$$

Rewriting Eq. 3 as a partial differential equation yields the following expression:

$$\bar{A} = A + \left(\frac{1}{2!V} \int_V w(d)d^2dV \right) \nabla^2 p \quad (4)$$

with

$$\frac{1}{V} \int_V w(d)dV = 1 \quad (5)$$

Therefore, the simplified nonlocal expression can be eventually expressed as follows:

$$\bar{A} = A + l^2 \nabla^2 A \quad (6)$$

In Eq. 6, the second-order gradient term is incorporated in addition to the original local counterpart. A length scale parameter l is used in order to weigh the gradient term, reflecting the characteristic length of the influence of the gradient.

Considering the nonuniform deformation of materials in micro- and nanoscales, the strain is usually used in order to characterize the material behaviors. By writing the nonlocal expression using the expression of strain and strain gradient, the strain gradient plasticity (SGP) theory can be written in the format of Eq. 6 as follows:

$$\bar{p} = p + l^2 \eta^2 \quad (7a)$$

where \bar{p} is the total accumulated plastic strain with its local counterpart p and the gradient term η , and l is the material length scale that characterizes the influence of the strain gradient.

Another approach was developed through a phenomenological theory of strain gradient plasticity based on gradients of rotation, which fits the framework of the couple stress theory (Fleck and Hutchinson 1993). By assuming that the strain energy density is only dependent on the second von Mises invariant of strain, the

strain gradient plasticity theory can be expressed based on that the strain energy density is only dependent on the overall effective strain as follows:

$$\bar{p} = \sqrt{p^2 + l^2 \eta^2} \quad (7b)$$

A more general expression (Voyiadjis and Abu Al-Rub 2005) for the strain gradient plasticity theory considering the expressions in Eqs. 6, 7a, and 7b was later proposed as follows:

$$\bar{p} = [p^\gamma + (l\eta)^\gamma]^{1/\gamma} \quad (8)$$

where γ is a fitting parameter.

Physically Based Material Length Scale

As shown in Eq. 8, the total accumulated plastic strain can be determined through its local counterpart and the strain gradient term. It becomes of great importance in the determination of the material length scale parameter, as it captures the intrinsic gradient effects for different materials. From the continuum theory, the flow stress can be written through the plastic strain given in Eq. 8 as follows:

$$\sigma = k[p^\gamma + (l\eta)^\gamma]^{1/m\gamma} \quad (9)$$

where m and k are material constants. Equation 8 is capable to predict the constitutive relations in both macroscopic scale through the local counterpart and microscopic scale through the coupling between the local and gradient terms. Therefore, by determining the length scale parameter, the strain gradient enhanced classical continuum theory is able to bridge the gap between large and small scales.

In Eq. 9, the local term p is related to the SSDs as it represents the uniform deformation at macroscopic scales. The plastic shear strain γ^p can be defined as a function of SSD density ρ_S as follows (Bammann and Aifantis 1982):

$$\gamma^p = b_S L_S \rho_S \quad (10)$$

where b_S is the magnitude of Burgers vector of SSDs and L_S is the mean space between SSDs. The local plastic strain ε_{ij}^p can then be determined from γ^p using Schmid orientation tensor M_{ij} as follows:

$$\varepsilon_{ij}^p = \gamma^p M_{ij} \quad (11)$$

where

$$p = \sqrt{2\varepsilon_{ij}^p \varepsilon_{ij}^p / 3} \quad (12)$$

When relating the plastic strain between macroscopic and microscopic scales, the Schmid orientation factor which is an average form of the Schmid tensor is always used (Bammann and Aifantis 1987; Dorgan and Voyiadjis 2003). With that being said, the local plastic strain can be expressed as follows:

$$p = b_S L_S \rho_S M \quad (13)$$

where M is the Schmid factor which is usually taken to be 0.5 (Bammann and Aifantis 1982).

The gradient term in Eq. 9 is related to the GNDs as it represents the nonuniform deformation that occurred in small scales. The gradient η can be expressed through GND density ρ_G and the Nye factor as follows (Arsenlis and Parks 1999):

$$\eta = \frac{\rho_G b_G}{\bar{r}} \quad (14)$$

where b_G is the magnitude of the Burgers factor of GNDs and \bar{r} is the Nye factor.

In microscopic scales, the flow stress can be related to the dislocation densities generated by the plastic deformation through Taylor's hardening law as follows:

$$\tau_S = \alpha_S G b_S \sqrt{\rho_S} \quad (15)$$

$$\tau_G = \alpha_G G b_G \sqrt{\rho_G} \quad (16)$$

where τ_S and τ_G are the shear flow stress corresponding to SSD density and GND density, respectively; b_S and b_G are the magnitudes of the Burgers vectors for SSDs and GNDs, respectively; G is the shear modulus; and α_S and α_G are statistical coefficients which account for the deviation from regular spatial arrangements of SSDs and GNDs populations, respectively. The total shear flow stress, τ_f , which is required to initiate a significant plastic deformation, can be obtained by coupling the flow stresses given by Eqs. 15 and 16 as follows:

$$\tau_f = \left[\tau_S^\beta + \tau_G^\beta \right]^{1/\beta} \quad (17)$$

where β is a constant fitting parameter.

A more general expression for the total shear flow stress can be written in terms of the total dislocation density from the combination of Eqs. 15, 16, and 17 as follows:

$$\tau = \alpha_S G b_S \sqrt{\rho_T} \quad (18)$$

with

$$\rho_T = \left[\rho_S^{\beta/2} + (\alpha_G^2 b_G^2 \rho_G / \alpha_S^2 b_S^2)^{\beta/2} \right]^{2/\beta} \quad (19)$$

The plastic flow stress is related to the shear flow stress through a constant Z as follows:

$$\sigma = ZT = \alpha_S Z G b_S \sqrt{\rho_T} \quad (20)$$

Comparing Eqs. 9 and 20 and considering Eqs. 13, 14, and 19, the following expressions can be determined:

$$\gamma = \beta/2 \quad (21)$$

$$m = 2 \quad (22)$$

$$k = \alpha_G G Z \sqrt{b_S / L_S M} \quad (23)$$

$$l = (\alpha_G / \alpha_S)^2 (b_G / b_S) L_S M \bar{r} \quad (24)$$

At the beginning, the length scales were determined as a constant value. However, it has been reported that L_S is not a constant but it is equal to the grain size initially and saturates toward values in the order of micrometer in large strains (Gracio 1994). L_S can be written as a function of the grain size and the plastic strain as follows:

$$L_S = \frac{Dd}{D + dp^{1/m}} \quad (25)$$

where d is the average grain size, D is the macroscopic characteristic size of the specimen, p is the equivalent plastic strain, and m is the hardening exponent. It can be seen from Eq. 24 that the length scale is proportional to L_S . Therefore, the length scales of crystalline materials are not constant during the deformation. This has also been experimentally and theoretically verified in the work of Voyiadjis and Abu Al-Rub (2005).

Due to the fact that the GND density during deformation is strain rate and temperature dependent, the Nye factor in Eq. 24 can be expressed as a function of the strain rate and temperature as follows (Voyiadjis and Almasri 2009):

$$\bar{r} = \frac{Ae^{(-E_r/RT)}}{1 + C\dot{p}^q} \quad (26)$$

where the temperature factor is incorporated through the Arrhenius equation; \dot{p} is the equivalent plastic strain rate; and A , C , and q are constants. Substituting Eqs. 25 and 26 into Eq. 24 yields a variable length scale that varies with the grain size, temperature, strain rate, and equivalent plastic strain as follows (Voyiadjis et al. 2011):

$$l = (\alpha_G/\alpha_S)^2 (b_G/b_S) M \left(\frac{\delta_1 \text{de}^{(E_r/RT)}}{(1 + \delta_2 p^{(1/m)}) (1 + \delta_3 (\dot{p})^q)} \right) \quad (27)$$

where δ_1 , δ_2 , and δ_3 are material parameters that need to be determined through experimental results. For single crystal materials, there is no effect of the grain size during the deformation. Therefore, the length scale for single crystal materials can be expressed without grain size d as follows:

$$l = (\alpha_G/\alpha_S)^2 (b_G/b_S) M \left(\frac{\delta_1 e^{(-E_r/RT)}}{(1 + \delta_2 p^{(1/m)}) (1 + \delta_3 (\dot{p})^q)} \right) \quad (28)$$

Determination of the Length Scales

Nanoindentation experiments are believed to be the most effective technique to determine the length scales. The indentations made using a conical or a Berkovich indenter have self-similar shapes, which means once the indent is initially formed, the size of the indent grows and the shape keeps similar. The GND and SSD densities can be calculated during the indentation process. The material hardness is related to the flow stress according to Tabor's factor. Therefore, the hardness can be written as a function of the dislocation densities through Eqs. 18, 19, and 20 as follows:

$$H = \kappa \sigma_f = Z \kappa \alpha_S G b_S \left[\rho_S^{\beta/2} + (\alpha_G b_G / \alpha_S b_S)^\beta \rho_G^{\beta/2} \right]^{1/\beta} \quad (29)$$

where κ is Tabor's factor relating the hardness and the flow stress. As shown in Eq. 29, the hardness is a function of both GND density and SSD density, which represents the deformation in the micro- and nanoscales where GNDs and SSDs interact with each other. However, in the case of large deformation, the effect of GNDs vanishes and the deformation is mainly attributed to the SSDs. The hardness in macroscopic scales, H_0 , is thus only related to the SSD density as follows:

$$H_0 = Z \kappa \tau_S = Z \kappa \alpha_S G b_S \sqrt{\rho_S} \quad (30)$$

Using Eqs. 29 and 30 and assuming that $\alpha_G = \alpha_S$ and $b_G = b_S$, the ratio $(H/H_0)^\beta$ can be derived as follows:

$$\left(\frac{H}{H_0} \right)^\beta = 1 + \left(\frac{\rho_G}{\rho_S} \right)^{\beta/2} \quad (31)$$

The nanoindentation experiments provide the material hardness as a function of the indentation depth. If the dislocation densities can be calculated as through the

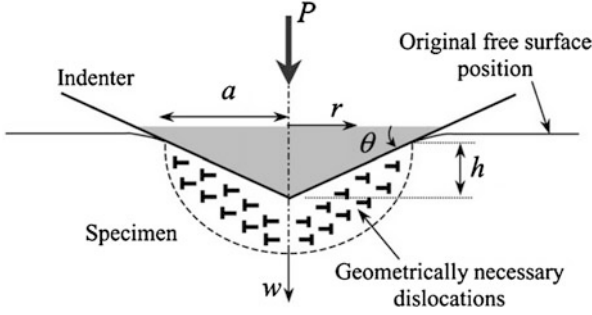


Fig. 1 Sample being indented by a conical indenter (Reprinted from Voyiadjis and Zhang (2015))

indentation depth, the hardness from the model can be compared with the hardness determined by the experiments.

The GND density during nanoindentation can be calculated through the geometry of the indenter. The cross-sectional profile of the indentation process using a conical indenter is shown in Fig. 1. The profile of a Berkovich indenter can be represented by the conical indenter if the surface angle θ is taken to be 0.358 (Radian), since the cross-sectional areas and the volumes are identical between the conical and pyramidal Berkovich indenters.

The dislocation loops are generated by the penetration of the indenter and therefore the plastic deformation volume can be assumed as a semisphere. The GND density can be determined as the total length of dislocations divided by the volume of the semisphere, where the dislocation length can be calculated based on the geometry of the indenter. Without considering the influence of the grain boundaries, the GND density can be written as follows (Nix and Gao 1998):

$$\rho_G = \frac{3 \tan^2 \theta}{2 b_G h} \quad (32)$$

It can be seen from Eq. 32 that the GND density decreases with the increasing indentation depth, proving that the nonuniform deformation decays with the increase of deformation and the nonlocal gradient becomes less significant.

In order to capture the hardening-softening phenomenon, grain boundaries were incorporated into the profile as shown in Fig. 2. When the plastic zone expands and reaches the grain boundaries, the diameter of the semisphere is equal to the average grain size.

However, a more accurate calculation can be approached by removing the volume occupied by the indenter from the semisphere (Yang and Vehoff 2007). The GND density can be written as follows (Voyiadjis and Peters 2010)

$$\rho_G = \frac{\frac{\pi}{b_G \tan \theta} h^2}{\frac{\pi}{12} d^3 - 8.19 h^3} \quad (33)$$

where d represents the average grain size in polycrystalline materials.

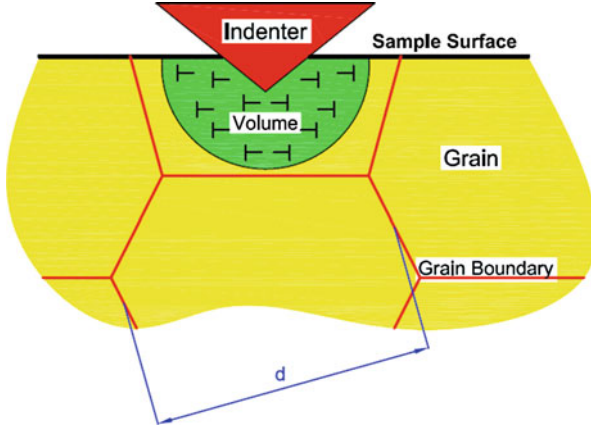


Fig. 2 Nanoindentation cross-sectional profile of a polycrystalline sample (Reprinted from Voyiadjis et al. (2011))

The SSD density can be derived from the Tabor's mapping of hardness-indentation depth from flow stress-plastic strain (Voyiadjis and Peters 2010). The SSD density ρ_S can be written as follows:

$$\rho_S = \frac{c\bar{r}\alpha_G^2 b_G \tan \theta}{lb_S^2 \alpha_S^2} \quad (34)$$

where c is a material constant of order 1 from Tabor's mapping.

Combining Eqs. 31, 32, and 34 and assuming that $\alpha_G = \alpha_S$ and $b_G = b_S$, the hardness of single crystal materials can be derived as a function of the indentation depth as follows:

$$H = H_0 \left\{ 1 + \left[\frac{3M \tan \theta}{2ch} \times \frac{\delta_1 e^{-E_r/RT}}{(1 + \delta_2 p^{1/m})(1 + \delta_3 \dot{p}^q)} \right]^{\beta/2} \right\}^{1/\beta} \quad (35)$$

For polycrystalline materials, similar derivation can be made through Eqs. 31, 33, and 34 as follows:

$$H = H_0 \left\{ 1 + \left[\frac{M \pi (h - h_1)^2}{\left(\frac{\pi}{12} d^3 - 8.19 h^3\right) c \tan^2 \theta} \times \frac{\delta_1 d e^{-E_r/RT}}{(1 + \delta_2 d p^{1/m})(1 + \delta_3 \dot{p}^q)} \right]^{\beta/2} \right\}^{1/\beta} \quad (36)$$

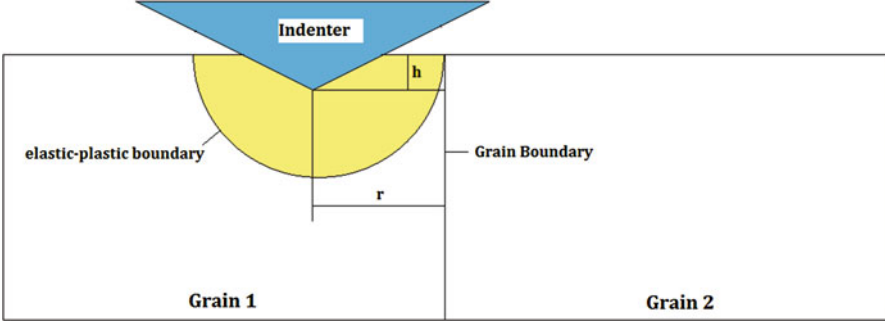


Fig. 3 Nanoindentation cross-sectional profile of a bicrystalline sample (Reprinted with permission from Voyiadjis and Zhang (2015)).

In order to isolate the influence of the grain boundary, bicrystalline materials are used and nanoindentations are made close to the single grain boundary near the grain boundary. The grain size in bicrystalline materials is not a parameter as it is in the size of the macroscopic scale. The distance r between the grain boundary and the indenter tip is used to represent the size of the plastic zone. When the plastic zone reaches the grain boundary, the radius of the semisphere is equal to the distance r as shown in Fig. 3.

By using the distance r instead of the grain size d , the GND density during nanoindentation of bicrystalline materials can be written as follows:

$$\rho_G = \frac{\frac{\pi}{b_G} \frac{h^2}{\tan \theta}}{\frac{2}{3}\pi r^3 - 8.19h^3} \quad (37)$$

Therefore, the hardness expression of a bicrystalline material during nanoindentation can be written as follows:

$$H = H_0 \left\{ 1 + \left[\frac{M\pi(h-h_1)^2}{\left(\frac{2\pi}{3}r^3 + 8.19h^3\right) c \tan^2 \theta} \times \frac{\delta_1 r e^{-E_r/RT}}{(1 + \delta_2 r p^{1/m})(1 + \delta_3 \dot{p}^q)} \right]^{\beta/2} \right\}^{1/\beta} \quad (38)$$

The length scale expression given by Eqs. 27 and 28 are incorporated into Eqs. 35, 36, and 38. The material parameters δ_1 , δ_2 , and δ_3 can be determined by comparing Eqs. 35, 36, and 38 with nanoindentation experimental results for single crystal, polycrystalline materials, and bicrystalline materials, respectively. Therefore, the length scale parameters can be determined for materials with different microstructures, respectively.

Furthermore, a better approach of the hardness expression can be applied through a cyclic plasticity model (Voyiadjis and Abu Al-Rub 2003) as follows:

$$H_{new} = H_{old} + C(h)H_{old} \quad (39)$$

with

$$C(h) = \sqrt{\frac{I}{h}} \quad (40)$$

where H_{old} is the hardness given by Eqs. 35, 36, and 38 and H_{new} is the corrected value using the cyclic model.

In order to write the hardness H as a function of the indentation depth h in Eqs. 35, 36, and 38, the equivalent plastic strain p is also required to be as a function of h . It is difficult to determine p as a function of h experimentally as the plastic deformation during nanoindentation is complicated. Therefore, finite element simulation is needed to capture this relationship (Voyiadjis and Peters 2010). Commercial finite element analysis software ABAQUS is used throughout an indentation problem. The testing sample is represented by a cube with dimension of 50 μm . The Berkovich indenter is modeled on top of the cubic sample as a blunt three-side pyramid as shown in Fig. 4. The tip of the indenter is set to be an equilateral triangle with 20 nm sides. The ABAQUS interaction module is used in order to simulate the contact between the indenter and the sample. A specific velocity field is assigned in order to drive the indenter to penetrate to the desired indentation depths according to the experiments. In order to show the different behaviors between FCC and BCC materials, a user material subroutine VUMAT is used during the simulation (Voyiadjis et al. 2011). After the indentation process is completed, the equivalent plastic strain p under the indenter can be viewed as shown in Fig. 5. The value of p at each time step can be determined through the contour plots as well as the indentation depth at each step. Therefore, the equivalent plastic strain as a function of the indentation depth can be determined by taking values of p and h at each step.

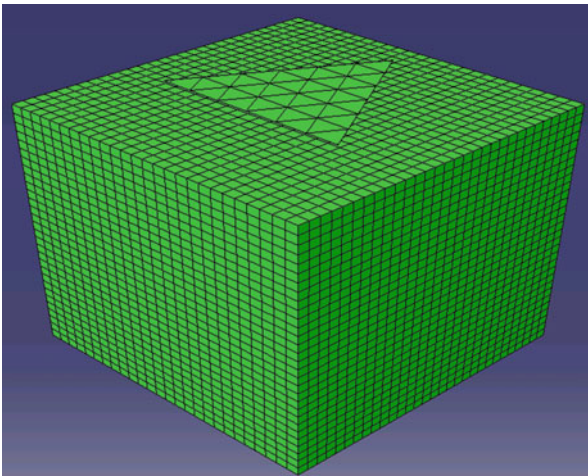


Fig. 4 Indentation model before the simulation (Reprinted from Voyiadjis and Zhang (2015))

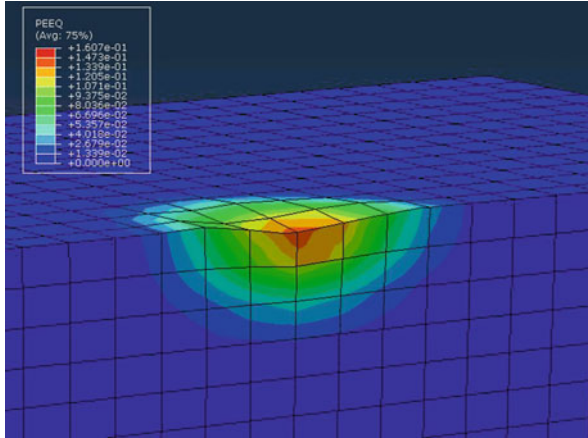


Fig. 5 Contour plot of equivalent plastic strain at the maximum indentation depth (Reprinted from Voyiadjis and Zhang (2015))

Applications on Single Crystal, Polycrystalline, and Bicrystalline Metals

Nanoindentation experiments are performed on metals with different microstructures in order to validate the prediction of the computational models and to determine the material parameters of the expressions of the length scales. Based on the temperature and strain rate indentation size effect (TRISE), nanoindentation experiments are conducted on single crystal and polycrystalline samples at different temperatures and strain rates. On bicrystalline samples, the experiments are carried out near the grain boundary at different distances in order to characterize the influence of the grain boundary on the material hardness. Nanoindentations at different strain rates are also performed at the same distance between the grain boundary and the indenter in order to confirm the rate dependency during nanoindentation on bicrystalline materials. The experimental results are compared with the prediction of hardness of the TRISE model. The length scales of materials with different microstructures are determined through the determinations of the material parameters from the comparisons. All testing samples are polished to acquire the accurate and consistent experimental results.

Sample Preparations

In nanoindentation, the hardness is not a direct measurement from the test. It is determined from the direct measurement of load and displacement through a tip area function (Oliver and Pharr 1992). The tip area function is determined through a model assuming that the indenter penetrates perpendicularly into a flat sample

surface. Therefore, in order to obtain accurate experimental results, the surface of the sample needs to be polished in order to approach the flat condition as assumed by the theory. It is required by the nanoindentation technique that the surface roughness must be smaller than one tenth of the maximum indentation depth. Therefore, if the hardness information at smaller indentation depths is needed, the surface roughness needs to be controlled to be lower.

Several polishing procedures are applied in order to improve the surface roughness. Mechanical polishing is usually applied firstly in order to level the entire surface to be horizontal. Silica carbide polishing papers with polishing particles of different sizes are used depending on the initial surface condition: the rougher the surface, the greater the size of polishing particles. After the use of each polishing paper, the surface is examined using a light microscope to make sure that the scratches on the surface are in the same size. Chemical-mechanical polishing is applied following the mechanical polishing when the polishing paper with the minimum size of polishing particles is used. In the chemical-mechanical polishing process, 50 nm colloidal silica or alumina polishing particles are used depending on the type of materials to be polished. By adjusting the PH values of the polishing slurries, chemical reactions between the polishing particles and the polishing sample can be initiated so that the surface roughness is lowered more effectively compared to the mechanical polishing process. The improvement of the surface quality of an iron sample is shown in Fig. 6. It shows that the surface roughness is improved and the defects such as voids and scratches are also improved.

In addition to the surface roughness, the plastic deformation layer on top of the surface is another concern for nanoindentation experiments. Due to the mechanical abrasion between the polishing particles and the sample, plastic deformation layer is induced no matter how small the polishing particles are used. The hardness varies if there is a plastic deformation that occurred in the indenter area, especially for tests in bicrystalline materials when the indentation depth is as small as a few tens of nanometers where the hardening-softening phenomenon is observed. Electro-polishing is widely used in order to remove the plastic deformation layer. Chemical reactions occur in the electrolyte solution and the peak material on the surface is removed by the electrical current during the electro-polishing process. As

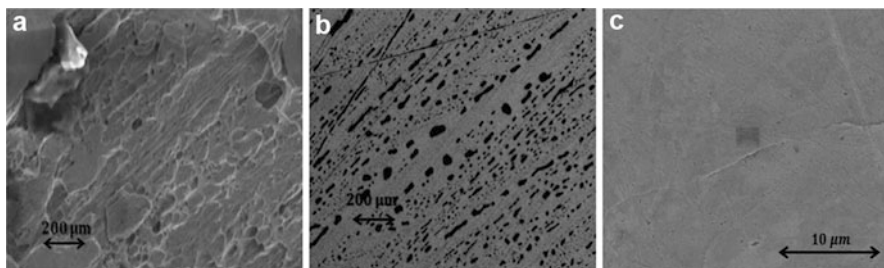


Fig. 6 SEM images of surfaces of testing sample in different polishing conditions: (a) without polishing, (b) after mechanical polishing, and (c) after chemical-mechanical polishing

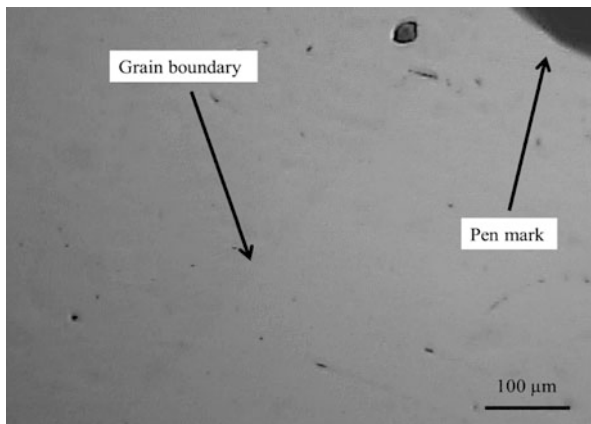


Fig. 7 Visualization of the grain boundary after vibratory on a bicrystalline aluminum (Reprinted from Voyiadjis and Zhang (2015))

no mechanical abrasion is induced, there is no plastic deformation created in the electro-polishing process. Although electro-polishing process is very effective in removing the plastic deformation, it has its disadvantages that the electrical current must be controlled carefully at a constant value; otherwise damages are induced on the surface. Vibratory polishing is another process that can remove the plastic deformation layer. In a vibratory polisher, a vertical vibration is added in addition to the rotation of the polishing pad. The down force pressure during polishing is thus minimized by the vibration so that only the peak material on the surface is removed in a very gentle manner. Due to the minimal down force pressure, the mechanical abrasion between the particles (usually alumina or colloidal silica) and the sample does induce significant plastic deformation. After the vibratory polishing process for bicrystalline samples, the grain boundary is observed under the light microscope as shown in Fig. 7. The pen mark is made on the surface of the bicrystal where the single grain boundary is formed during the growth of the bicrystal.

Temperature and Strain Rate Dependency on Single Crystal and Polycrystalline Metals

The TRISE model incorporates the temperature and strain rate parameters into the expression of the length scale. As the expression of hardness is written using the length scale, the hardness is predicted to be dependent on the temperature and the strain rate. In order to verify the prediction of the TRISE model and material parameters needed for the length scales, nanoindentation experiments are conducted on various metals with single crystal and polycrystalline microstructures at different temperatures and strain rates. The material parameters are determined by comparing the TRISE model and the experiments. The length scales are eventually

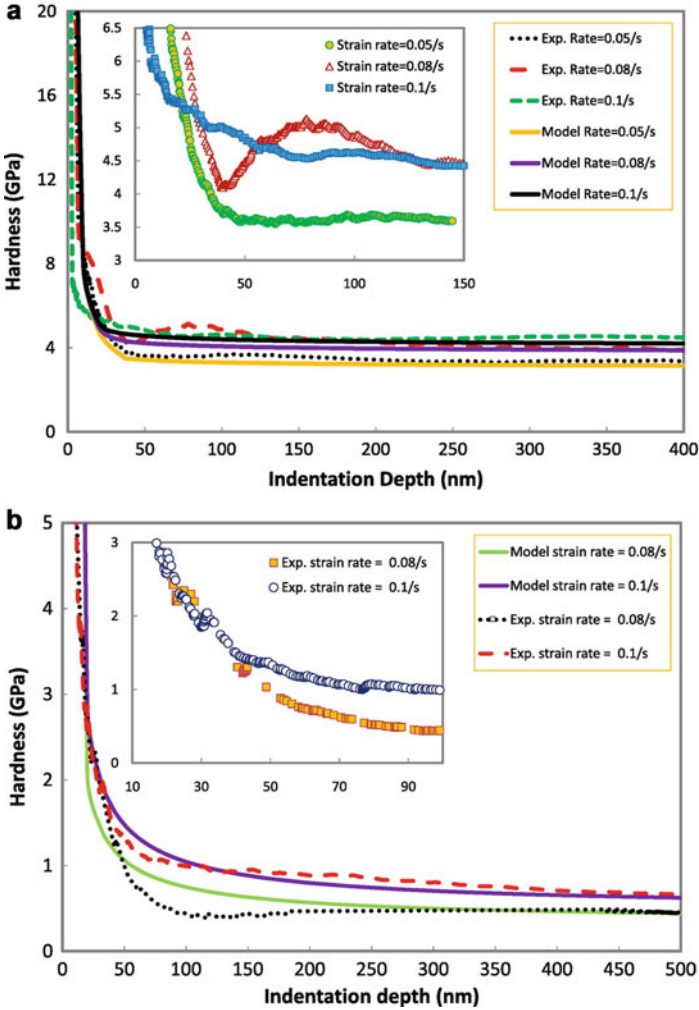


Fig. 8 Comparisons between TRISE models and nanoindentation experiments: (a) copper single crystal and (b) aluminum single crystal (Reprinted from Voyiadjis et al. (2011))

determined for single crystal and polycrystalline materials at different strain rates and temperatures.

In aluminum and copper single crystals, nanoindentation experiments are performed on a polished surface at different strain rates of 0.05 s^{-1} , 0.08 s^{-1} , and 0.10 s^{-1} . It shows from the experimental results in Fig. 8 that the hardness increases with the increasing strain rate for both metals. The TRISE model for single crystal given by Eq. 35 is applied by giving different values for the parameter equivalent plastic strain rate $\dot{\bar{p}}$. It shows a good agreement between the experimental results and the predictions from the hardness expression. This proves that the TRISE

Table 1 Materials parameters used in Eq. 35 for single crystals (Reprinted from Voyiadjis et al. 2011)

Parameters	h_1 (nm)	δ_1	δ_2	δ_3	M	c	θ (Rad)	m	β	q
Copper	32	7,050	50	0.01	0.5	1	0.358	0.474	2	0.3
Aluminum	50	4,000	48	0.01	0.5	1	0.358	0.474	1.5	0.3

model is able to predict the influence given by the strain rates. Moreover, there is no hardening-softening phenomenon observed which validates the assumption that no dislocation is accumulated with the absence of grain boundaries.

The material parameters shown in Table 1 are obtained through the comparisons and the length scales of copper and aluminum single crystals as shown in Fig. 9.

Although no hardening-softening phenomenon is observed due to the absence of the grain boundaries, the experimental results and the TRISE model both show that the hardness decreases with increasing indentation depth. This is due to the generation of GNDs at small depth where nonuniform deformation is significant during the initial formation of the indent. The GNDs also interact with the SSDs as obstacles of the movement of SSDs, giving rise to the increase of hardness at smaller depths. The material length scales for both single crystals show the length scale decrease with the increase of the equivalent plastic strain, which means the strain gradient is greater when deformation is smaller. The determination of the length scale verifies the strain gradient plasticity theory that the stress mechanical responses are greater when the strain gradients are higher. It also shows in the length scales that a variable length scale expression needs to be used instead of the constant length scale determined in the previous research, which proves that a variable length scale is more realistic and the expression can be applied to different problems.

Nanoindentation experiments are conducted on polycrystalline copper and aluminum in order to capture the hardening-softening phenomenon as well as the strain rate dependency. TRISE model given by Eq. 36 is applied to compare with the experimental results. As shown in Fig. 10, the hardness initially decreases with indentation depth at very small depth less than 50 nm. The hardening-softening phenomenon is observed following the decreasing hardness segment. The experimental results show that at very small indentation depth where the plastic deformation does not reach the grain boundaries, the polycrystalline materials behave similarly with respect to single crystal materials. However, due to the presence of the grain boundaries, the expansion of the plastic zone is constrained and dislocations start to accumulate near the grain boundary. The local increase of the dislocation density gives rise to the increase of the material hardness, causing the hardening-softening phenomenon. Nanoindentation experiments are also performed at different strain rates. It shows in both experiments and TRISE model that the hardness increases with the increasing strain rate, verifying the strain rate dependency on polycrystalline materials.

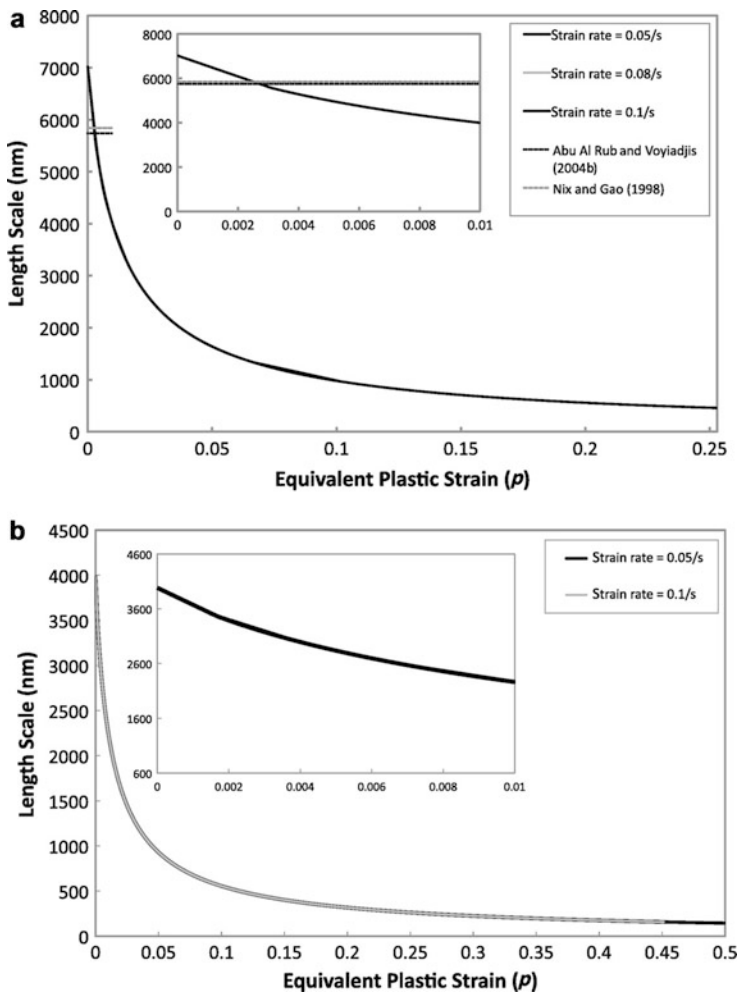


Fig. 9 Material length scales of single crystals: (a) copper and (b) aluminum (Reprinted from Voyiadjis et al. (2011))

The length scales for polycrystalline copper and aluminum can be determined by comparing the experimental results and TRISE model in determining the material parameters. Using the parameters given in Table 2, the length scales of both materials are obtained as a function of the equivalent plastic strain as shown in Fig. 11.

As shown in Fig. 11, the length scale decreases with the increasing strain rate. When the strain rate is increased, the deformation occurs faster and the dislocations do not have the sufficient time to generate and are trapped by the local nonuniform deformation. This causes the additional increase of the strain gradient, resulting in the increase of the material hardness.

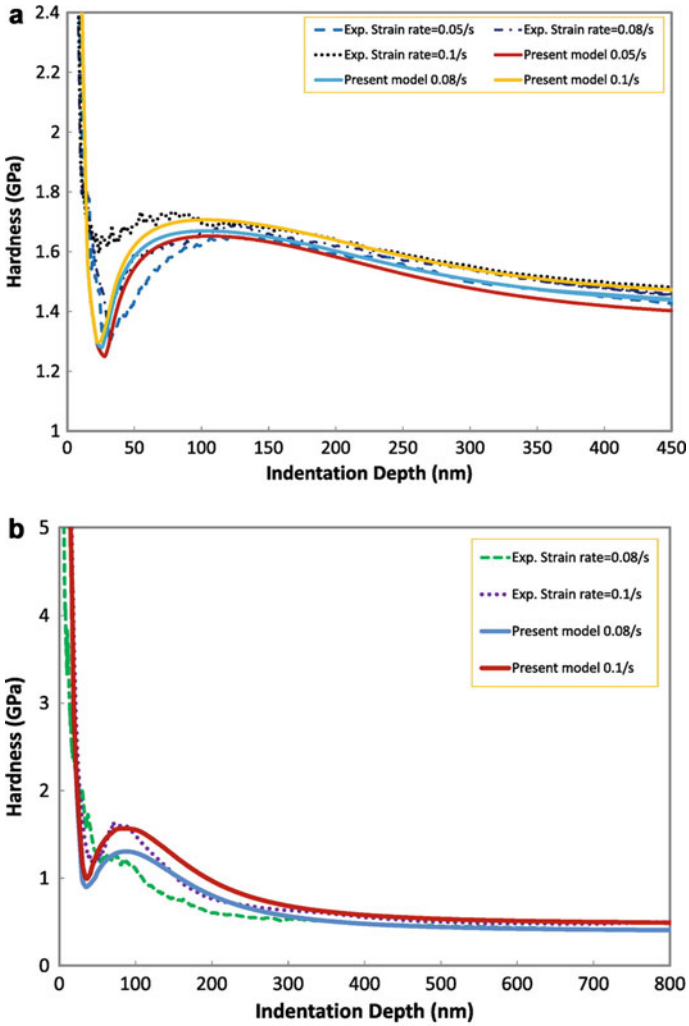


Fig. 10 Nanoindentation experiments and TRISE model of polycrystalline materials: (a) copper and (b) aluminum at different strain rates (Reprinted from Voyiadjis et al. (2011))

Table 2 Materials parameters used in Eq. 36 for polycrystalline materials (Reprinted from Voyiadjis et al. 2011)

Parameters	h_1 (nm)	δ_1	δ_2	δ_3	d (nm)	M	c	θ (Rad)	m	β	q
Copper	26	2.1	1	0.01	1,000	0.5	1	0.358	0.474	2	0.3
Aluminum	41	1.9	1	0.01	1,000	0.5	1	0.358	0.474	1.5	0.3

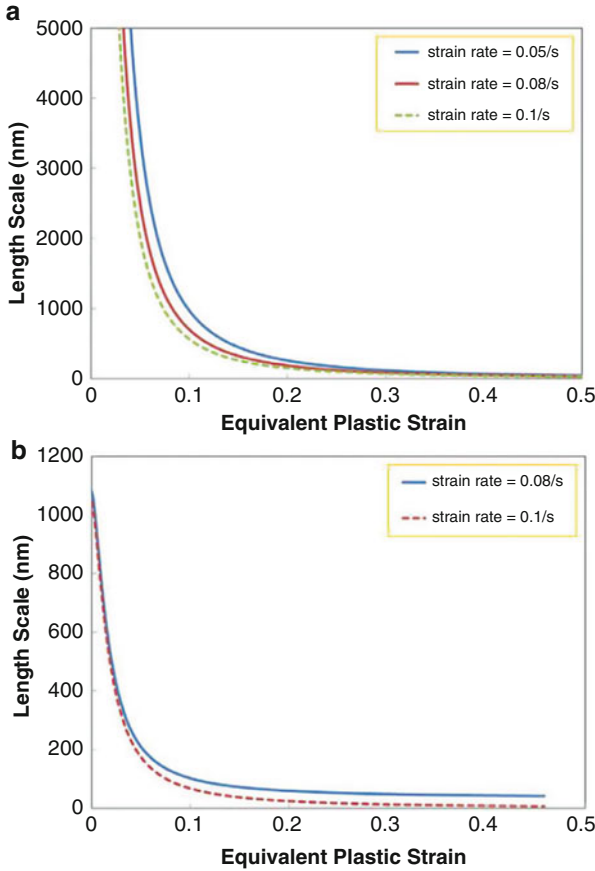


Fig. 11 Length scales as a function of the equivalent plastic strain for polycrystalline materials: (a) copper and (b) aluminum at different strain rates (Reprinted from Voyiadjis et al. 2011)

Nanoindentation experiments are performed on iron (Bahr et al. 1999) and gold (Volinsky et al. 2004) at different temperatures in order to study the temperature dependency. Different temperatures are given to the TRISE model and the comparisons between the experiments and models are presented in Fig. 12 Voyiadjis and Faghihi (2012). It shows that the hardness decreases with the increase of the temperature. The higher mobility of the dislocations causes the decrease of the dislocation density and thus the hardness decreases.

The length scales of iron and gold are obtained through the comparison and it shows in Fig.13 that the length scales increase with the increasing temperature, causing the decrease in the material hardness. As iron is a BCC metal and gold is an FCC metal, it shows that the TRISE model works for both BCC and FCC metals in predicting the indentation size effect and determining the length scales.

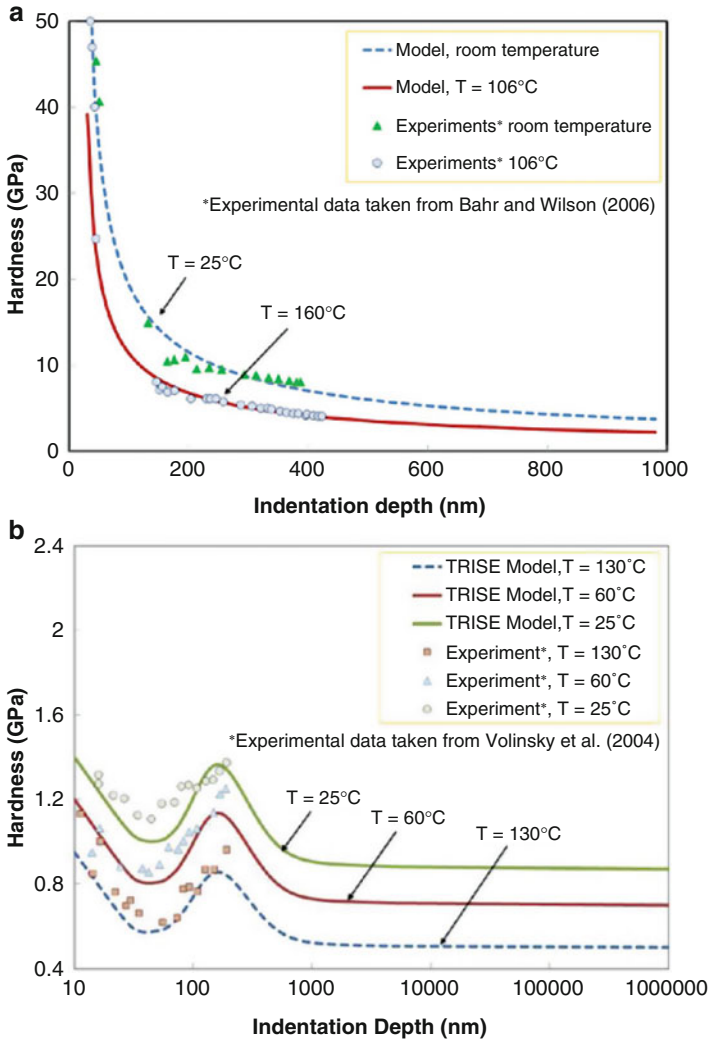


Fig. 12 Nanoindentation experiments at different temperatures on (a) iron single crystal and (b) gold polycrystalline film (Reprinted from Voyiadjis and Faghihi (2012))

Influence of the Grain Boundaries on Bicrystal Metals

As discussed in the previous section, the distance between the indenter and the grain boundary during nanoindentation has an impact on the hardness behavior of materials with grain boundaries. In the case of polycrystalline materials, the grain size d characterizes the interactions between the dislocations and the grain boundaries. The distance between the indenter and the grain boundary becomes

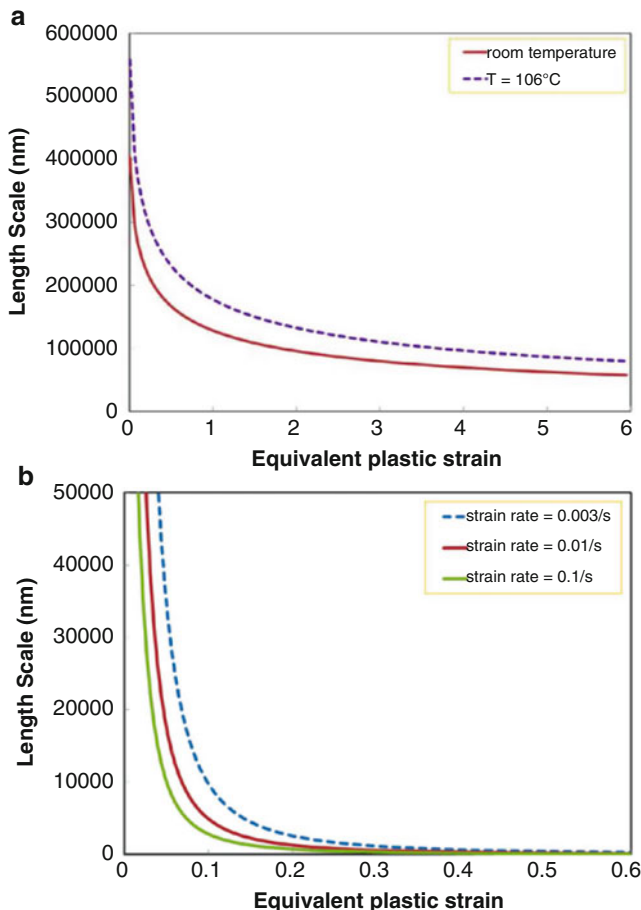


Fig. 13 Length scales at different temperatures: (a) iron and (b) gold (Reprinted from Voyiadjis and Faghihi (2012))

greater when the grain size increases. Nanoindentation experiments are performed on Nickel samples with different grain sizes (Yang and Vehoff 2007). The TRISE model given by Eq. 36 is applied by assigning different values to the grain size d (Voyiadjis and Faghihi 2012). The comparison between the experiments and the prediction of TRISE model is shown in Fig. 14. It shows in the experiments that as the grain size increases, the hardness during the hardening-softening segment decreases. As the dislocations only accumulate when they move to the grain boundaries, the greater grain size allows the dislocations to move in a greater space comparing to the condition of smaller grain size. The dislocation density thus decreases, resulting in the decrease in hardness. The experimental results show a good agreement with the predictions given by the TRISE model at different grain sizes as shown in Fig. 14.

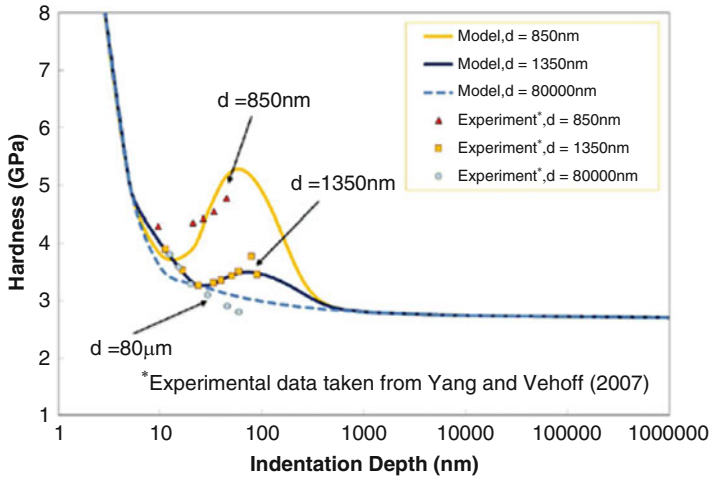


Fig. 14 Comparison between the experiments and the TRISE model at different strain rates (Reprinted from Voyiadjis and Faghihi (2012))

The size of the plastic zone increases with the increasing indentation depth. When the grain size is smaller, the plastic zone reaches the grain boundary at lower depth. As the grain size increases, the plastic zone reaches the grain boundary at greater depth. When the grain size is as much as $80 \mu\text{m}$, the size of the plastic zone is smaller than the grain size. There is no accumulation of the grain boundary and no hardening-softening phenomenon is observed. The length scales at different grain sizes are obtained from the comparison. It shows in Fig. 15 that the length scale decreases as the grain size decreases. The grain boundaries act as constrains that block the influence of the strain gradient, which causes the smaller characteristic length scale when the grain size is smaller.

In order to isolate the investigation on the influence of the grain boundary, bicrystalline materials are tested near the single grain boundary at different distances. After the Aluminum sample is polished, the single grain boundary is observed as shown in Fig. 7. However, it is only visible using a $10\times$ microscope in the Nanoindenter. The distance between the indenter and the grain boundary can be only measured using a $40\times$ microscope. In order to identify the position of the grain boundary, the two indents are made on two points on the grain boundary under a $10\times$ microscope. The grain boundary can thus be represented by the straight line connecting the two marking indents under the $40\times$ microscope. A straight line of nanoindentations are performed near the grain boundary with the angle between the line of indents and the grain boundary to be 20° as shown in Fig. 16a (Voyiadjis and Zhang 2015).

It is noted that the depths of the indentations are 500 nm and the spacing is $5 \mu\text{m}$. In order to make sure that there is no interaction between the close indents, experiments are made on other locations and the results show that there is no

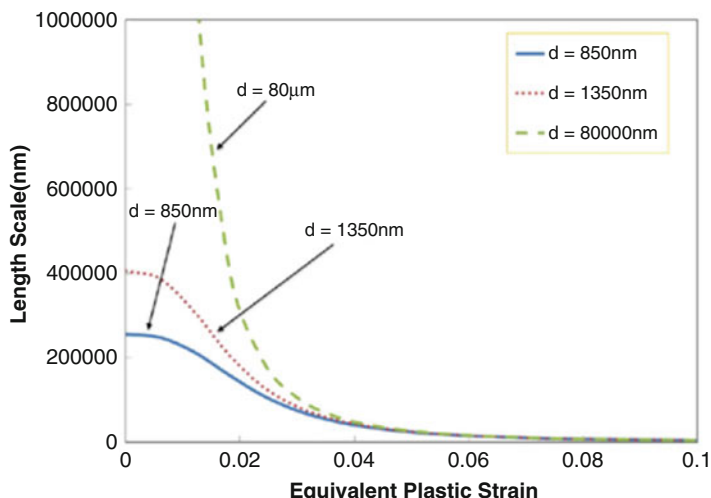


Fig. 15 Length scales as a function of the equivalent plastic strain at different grain sizes (Reprinted from Voyiadjis and Faghihi (2012))

significant interaction observed. The difference in the microstructures from two grains is not considered here. Therefore, the five indents on one side of the grain boundary are selected for the investigation. The distances between the center of the indents and the grain boundary are measured using the scale in the image as shown in Fig. 16b. The experimental results of the five indents are presented in Fig. 17. It shows that for the two indents that are far from the grain boundary, there is no hardening-softening phenomenon observed. The grain boundary does not have the impact on the accumulation of dislocations when the distance is large, showing the single crystal behavior. It also shows in Fig. 17 that hardening-softening phenomenon is observed for the rest three indents and as the distance becomes smaller, there is a greater hardening effect. This observation verifies the assumption of the influence of the grain boundary that more dislocations are accumulated between the indenter and the grain boundary when the distance is smaller. The higher dislocation density at smaller distances causes the increase in hardness.

The developed TRISE model for bicrystalline materials given by Eq. 38 is applied by assigning different values of the distance r measured from the experiments. The comparison between the TRISE model and the experimental results of the three indents with the grain boundary effect is given in Fig. 18. The developed TRISE model shows its capability to predict the hardness with the presence of the single grain boundary. It also reflects the prediction that there is a higher hardening effect as the indents are made closer to the grain boundary. The material parameters in Eq. 38 are determined by the comparison between the TRISE model and the experimental results. The material length scales are obtained as shown in Fig. 19 through the determination of material parameters given in Table 3.

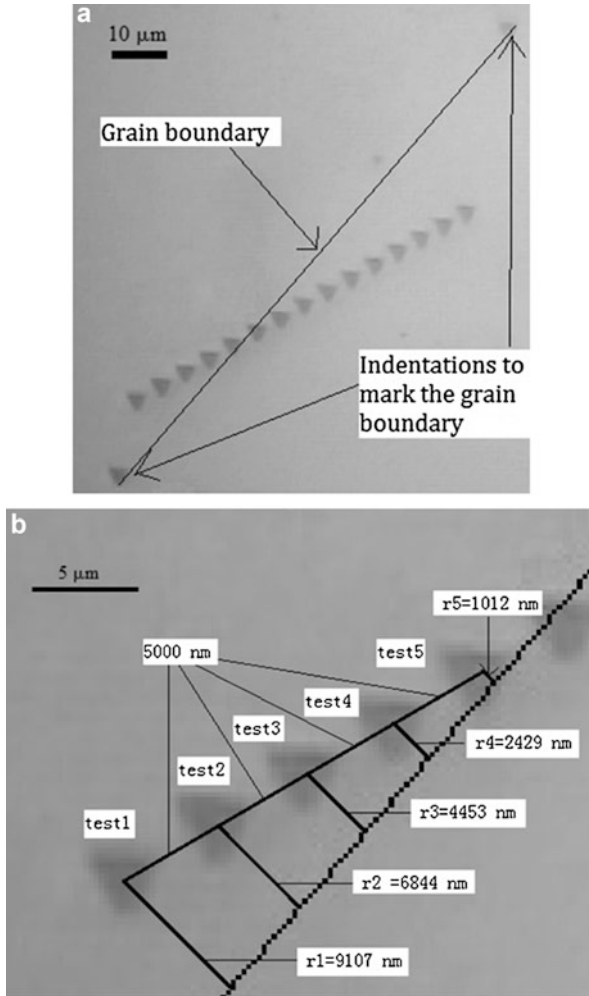


Fig. 16 Nanoindentations near the grain boundary at different distances: (a) image under the 40× microscope and (b) the magnified image of (a) (Reprinted from Voyiadjis and Zhang (2015))

The length scale decreases as the distance between the indenter and the grain boundary decreases. The grain boundary and the indenter both act as obstacles of the influence of the strain field. Therefore, the impact range of the strain gradient is shorter when the distance becomes smaller, resulting in a smaller characteristic length scale. The decrease in the length scale, in return, causes the increase of the strain gradient and thus causes the additional increase on material hardness.

Similar nanoindentation experiments are performed on a bicrystalline copper which is also an FCC metal. The grain boundary is also represented by a straight line connecting the two marking indents made on two points of the straight

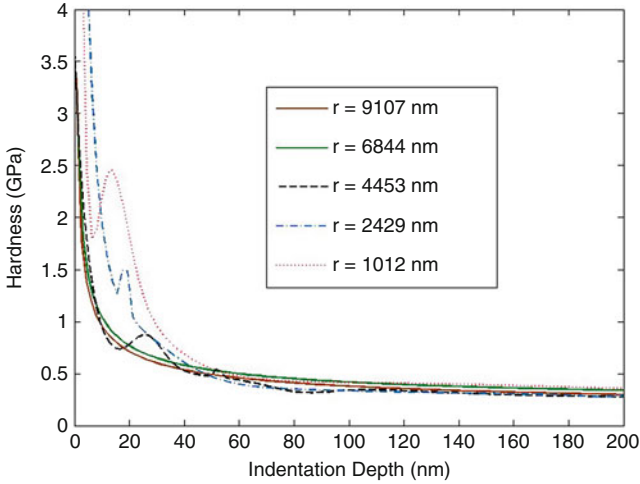


Fig. 17 Hardness versus indentation depth curves for five indents with different distances between the indenter and the grain boundary. The two *solid curves* represent the indents without hardening-softening phenomenon. The three *dashed lines* show the influence of the grain boundary on material hardness (Reprinted from Voyiadjis and Zhang (2015))

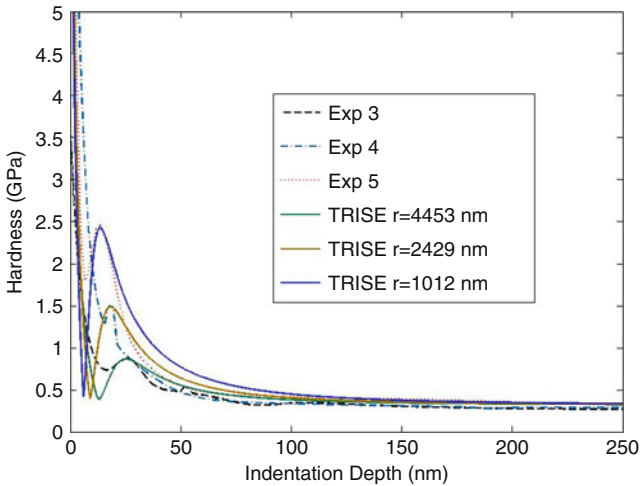


Fig. 18 Comparison of TRISE model with experimental results for the three indents with hardening-softening effect. The *solid lines* are obtained from the TRISE model and the *dashed lines* are the experimental results (Reprinted from Voyiadjis and Zhang (2015))

grain boundary. A line of indents are made across the grain boundary with the angle between indentation line and the grain boundary to be 5° as shown in Fig. 20a (Zhang and Voyiadjis 2016). The image in Fig. 20a is zoomed in order to show the details of the distance between the indent and the grain boundary as

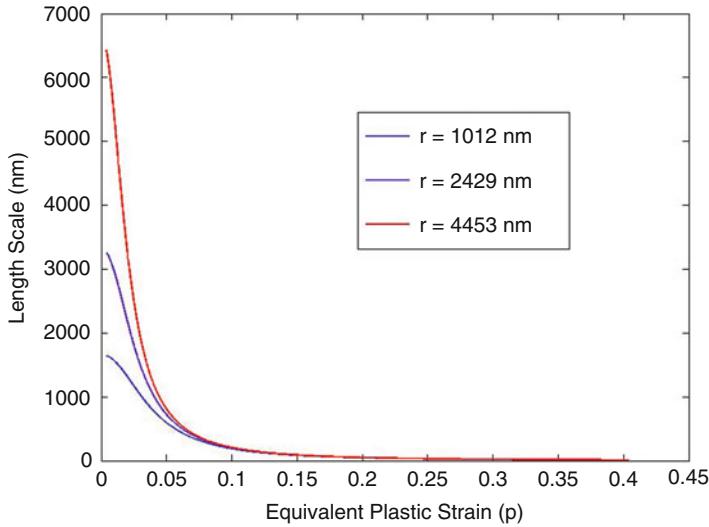


Fig. 19 Length scales as a function of the equivalent plastic strain at different values of r (Reprinted from Voyiadjis and Zhang (2015))

Table 3 Material parameters for bicrystalline Aluminum (Reprinted from Voyiadjis and Zhang 2015)

Material parameters	h_1 (nm)	δ_1	δ_2	δ_3
Aluminum	10	1.9	135	0.01

shown in Fig. 20b. The distances are measured using the scale considering the magnification. Artificial effect is added to Fig. 20b in order to show the clear contrast between the indents and the sample surface.

The experimental results of the three indents shown in Fig. 20b on the same side of the grain boundary as well as the indent made right on the grain boundary are presented in Fig. 21. The indent on the grain boundary solely decreases with the increasing indentation depth. When the indent is made on the grain boundary, the dislocations start to generate into the two grains on both sides of the grain boundary. There is no additional obstacle of the generation of dislocations and therefore no hardening-softening phenomenon is observed. This proves the assumption of the influence of the grain boundary on the other side. Hardening-softening effect is observed for the three indents within a close distance to the grain boundary. It also shows in Fig. 21 that as the distance becomes smaller, there is a greater hardening effect, similarly to the observation made from the experiments of Aluminum bicrystal. The experiments on Copper bicrystal confirm the influence of the grain boundary on the material hardness on FCC metal.

Due to the tip rounding of the indenter, strain hardening may be induced during nanoindentation at very small indentation depths. In order to confirm that the hardening is solely attributed to the accumulation of dislocations, the elastic moduli

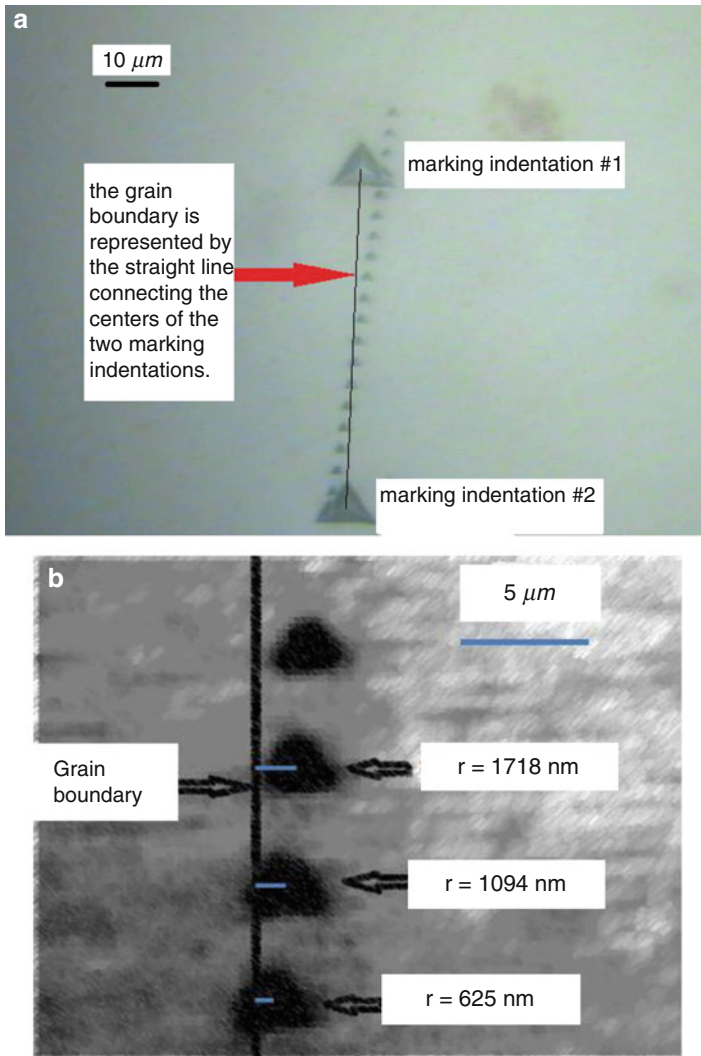


Fig. 20 Nanoindentation experiments with different distances from the grain boundary: (a) image under the 40× microscope and (b) a zoomed view of (a) (Reprinted from Zhang and Voyiadjis (2016))

of the three indents at different distances are presented in Fig. 22. It shows that after the depth of 20 nm, the elastic moduli from the three indents are constants on the average. This means that after the depth of 20 nm where the hardening effect is captured, there is no strain hardening effect induced as the elastic modulus is not dependent on the indentation size effect but only dependent on the strain hardening. The information of elastic moduli confirms that the hardening effect observed for

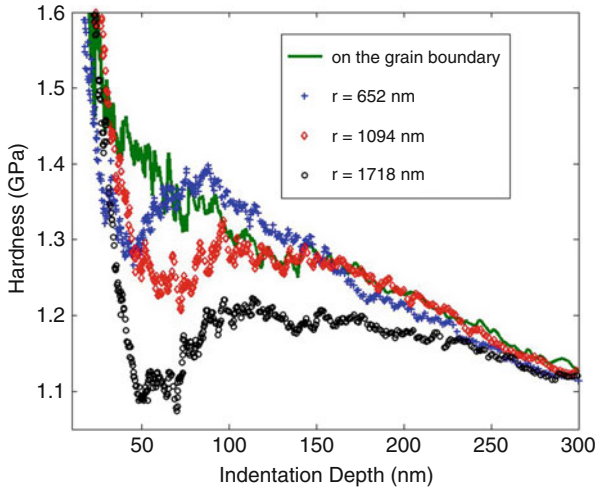


Fig. 21 Hardness versus indentation depth curves from nanoindentation experiments at different distances to the grain boundary and *right* on the grain boundary (Reprinted from Zhang and Voyiadjis (2016))

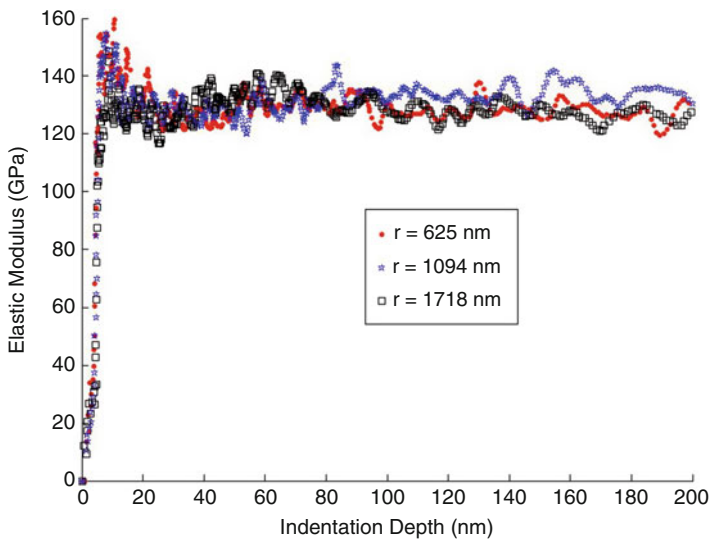


Fig. 22 Elastic modulus as a function of the indentation depth for three indents near the grain boundary (Reprinted from Zhang and Voyiadjis (2016))

the Copper bicrystal is only attributed to the accumulation of dislocations between the indenter and the grain boundary.

The developed TRISE model given by Eq. 38 is applied for the copper bicrystal with different values of distances. The comparison between the TRISE model and

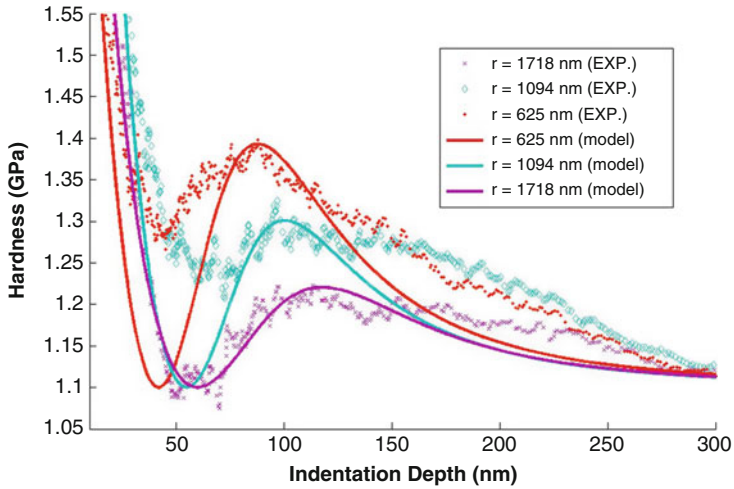


Fig. 23 Comparison between the TRISE model and nanoindentation experiments at different distances between the indenter and the grain boundary (Reprinted from Zhang and Voyiadjis (2016))

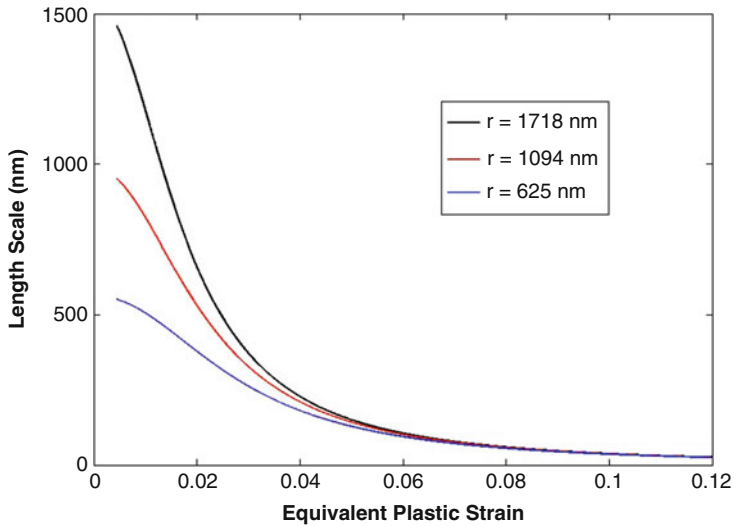
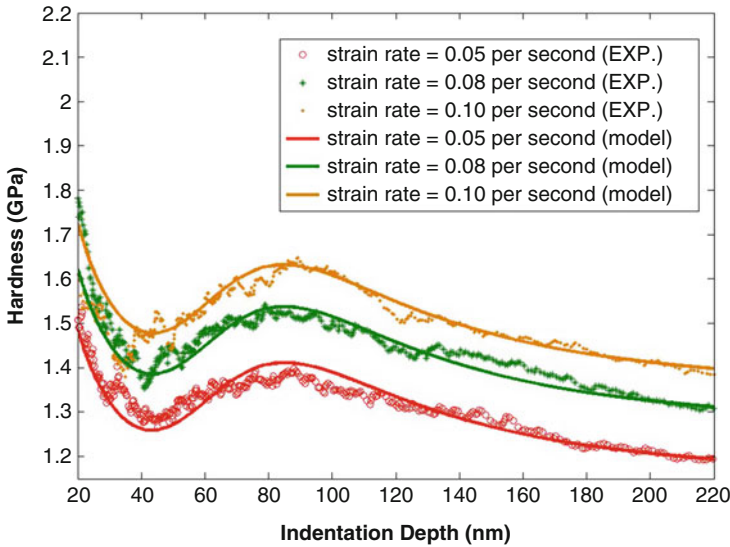


Fig. 24 Material length scales as a function of the equivalent plastic strain for Copper bicrystal at different distances between the indenter and the grain boundary (Reprinted from Zhang and Voyiadjis (2016))

Table 4 Material parameters of Copper bicrystal (Reprinted with permission from Zhang and Voyiadjis 2016)

Material parameters	h_1 (nm)	δ_1	δ_2	δ_2
Copper	42	2.0	3.0	0.01

**Fig. 25** Hardness vs. indentation depth for nanoindentation experiments at different strain rates with the same distance r (Reprinted from Zhang and Voyiadjis (2016))

experimental results is presented in Fig. 23. The length scales at different distances for the Copper bicrystal are determined as shown in Fig. 24 using the material parameters determined from the comparison as shown in Table 4.

The comparison shows the capability of the developed TRISE model on Copper bicrystal. As shown in Fig. 24, the length scale decreases as the distance r becomes smaller, which confirms the influence of the grain boundary on the material behavior of FCC metals.

Moreover, the strain rate dependency is studied for Copper bicrystal. Nanoindentation experiments are performed at the same distance from the grain boundary at different strain rates of $0.05s^{-1}$, $0.08s^{-1}$, and $0.10s^{-1}$. As shown in Fig. 25, the hardness increases with the increasing strain rates. It is worth noting that this increase is not only for the hardening-softening segment but for the entire depth. This is because the strain rate has the influence on material behaviors in both micro- and macroscales.

The TRISE model given by Eq. 38 is applied with different values of strain rates. The material length scales are determined using different values of strain rates and the material parameters given by Table 4. It shows in Fig. 26 that the length scale

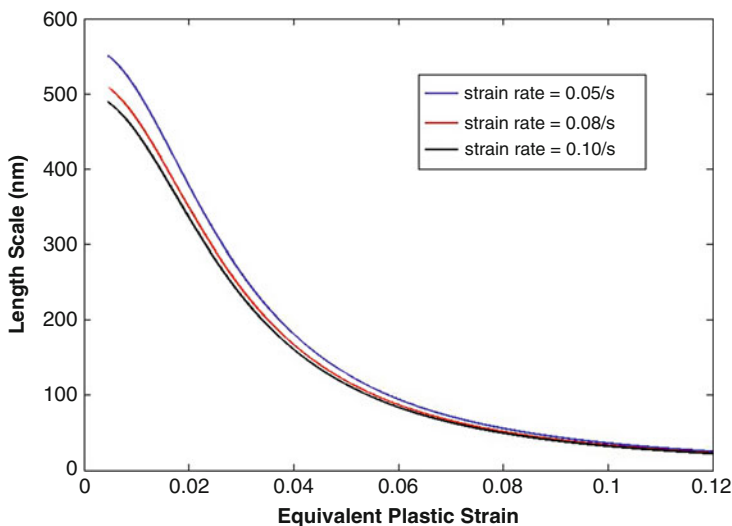


Fig. 26 Length scales as a function of the equivalent plastic strain at different strain rates for Copper bicrystal (Reprinted from Zhang and Voyiadjis (2016))

decreases with the increase of the strain rate. The strain rate dependency for Copper bicrystal is similar with the observation made from the polycrystalline copper.

Summary and Conclusion

The size effect during nanoindentation is addressed in this chapter. The classical continuum theory is not capable in predicting the indentation size effect as it does not incorporate the length scale in the constitutive expression. Strain gradient plasticity theory is applied in order to capture the size effect. The expression of a physically based length scale is determined from the strain gradient plasticity theory. In order to determine the material length scales, nanoindentation experiments are performed on single crystals, polycrystalline materials and bicrystalline materials. The constitutive equation is mapped into hardness versus indentation depth and the material parameters are determined from the expression of hardness and the experimental results.

The material length scale is dependent on the temperature, equivalent plastic strain rate, and grain size in polycrystalline materials. A TRISE model is applied in order to capture the dependencies of temperature, strain rate, and grain size. Nanoindentation experiments are performed on single crystal and polycrystalline materials. From the comparison between the TRISE model and experimental results, the length scales at different temperatures, strain rates, and grain sizes are determined.

The results of nanoindentation experiments on single crystal Copper and Aluminum show that the hardness solely decreases with the increase of the indentation depth. The TRISE model for single crystal is able to predict the hardness obtained from experiments of single crystal materials. The nanoindentation experiments on polycrystalline Copper and Aluminum show the hardening-softening effect due to the presence of grain boundaries. The TRISE model is capable in predicting the hardening-softening phenomenon. The temperature dependency is addressed by the TRISE model at different temperatures and nanoindentation experiments on iron and gold. Both TRISE model and experiments show that the hardness decreases with the increasing temperatures. Nanoindentation experiments are performed at different strain rates on both polycrystalline and single copper and aluminum. It is shown by the TRISE model and experimental results that hardness increases as the strain rate increases. The material length scales are determined by the comparison between the TRISE model and the experiments. The dependencies of length scales on the temperature, strain rate, and grain size verify the theory that higher strain gradient causes greater material hardness.

In order to isolate the influence of the grain boundary, bicrystalline copper and aluminum are used and nanoindentation experiments are performed near the grain boundary at different distances. The TRISE model is developed based on the structure of bicrystals. The length scales are determined by comparing the developed TRISE model and the experimental results on bicrystalline materials.

The experimental results on bicrystalline copper and aluminum show that the material behaves like a single crystal when the indents are made with large distances from the grain boundary. The hardening-softening phenomenon is only observed for the indents made in close proximity to the grain boundary. The experimental observation validates that the increase of hardness is attributed to the presence of the grain boundary. The accumulation of dislocations near the grain boundary causes the increase of dislocation density, resulting in the additional hardening effect. The influence of the grain boundary is further investigated through nanoindentation experiments near the grain boundary at different distances. The TRISE model is developed for bicrystalline materials by replacing the grain size d in polycrystalline models with the distance r between the indents and the grain boundary. The developed TRISE model shows its capability to predict the fact that there is a stronger hardening effect when the distance r becomes smaller as shown by nanoindentation experiments, providing a new type of size effect with respect to the distance r . The length scales of bicrystalline copper and aluminum show that as the indents are made closer to the grain boundary, the length scales decrease. The grain boundary and the indenter act as obstacles that prevent the strain gradient from influencing the space outside the volume constrained by the obstacles. Therefore, the characteristic length scale at smaller distance is lower and in return causes the increase in the strain gradient and dislocation density. The increase of dislocation density causes the additional hardening effect during nanoindentation. The rate dependency is also investigated on bicrystalline copper through nanoindentation experiments at different strain rates with a fixed distance from the grain boundary. Both experimental results and the TRISE model show an increase of hardness with

the increasing strain rate, similarly to the rate dependency in single crystal and polycrystalline materials.

For all the length scales determined in this chapter, it shows that the length scale decreases with the increasing equivalent plastic strain and approaches to zero when the deformation is large. This behavior of the length scale proves the strain gradient plasticity theory in addressing the size effect during nanoindentation. When the indentation depth is small, the shape of the indent is initially formed. GNDs are required to accommodate the nonuniform deformation in the formation of the indent. As the indenter penetrates deeper, the indent grows to a greater size with a similar shape because of the use of the self-similar Berkovich indenter. The amount of uniform deformation increases and SSDs are required to generate the uniform deformation. Therefore, the GND density decreases gradually with the increase of the indentation depth and the strain gradient becomes smaller. The influence of the decreasing strain gradient causes the decrease of the characteristic length. When the deformation becomes larger in the macroscopic scale, uniform deformation dominates and there is no influence of the strain gradient, resulting in a zero value of the length scale in large deformations.

References

- E.C. Aifantis, *Int. J. Eng. Sci.* **30**, 10 (1992)
A. Arsenlis, D.M. Parks, *Acta Mater.* **47**, 5 (1999)
D.F. Bahr, D.E. Wilson, D.A. Crowson, *J. Mater. Res.* **14**, 6 (1999)
D.J. Bammann, E.C. Aifantis, *Acta Mech.* **45**, 1–2 (1982)
D.J. Bammann, E.C. Aifantis, *Mater. Sci. Eng.: A* **309** (1987)
M.R. Begley, J.W. Hutchinson, *J. Mech. Phys. Solids* **46**, 10 (1998)
X. Chen, N. Ogasawara, M. Zhao, N. Chiba, *J. Mech. Phys. Solids* **55**, 8 (2007)
M.S. De Guzman, G. Neubauer, P. Flinn, W.D. Nix, *MRS Proc* **308**, 613 (1993)
R.J. Dorgan, G.Z. Voyiadjis, *Mech. Mater.* **35**, 8 (2003)
A.C. Eringen, D.G.B. Edelen, *Int. J. Eng. Sci.* **10**, 3 (1972)
N.A. Fleck, J.W. Hutchinson, *J. Mech. Phys. Solids* **41**, 12 (1993)
N.A. Fleck, J.W. Hutchinson, *Adv. Appl. Mech.* **33** (1997)
N.A. Fleck, G.M. Muller, M.F. Ashby, J.W. Hutchinson, *Acta Metall. Mater.* **42**, 2 (1994)
J.J. Gracio, *Scr. Metall. Mater.* **31**, 4 (1994)
E. Kroner, *Int. J. Appl. Phys.* **40**, 9 (1967)
Q. Ma, D.R. Clarke, *J. Mater. Res.* **10**, 4 (1995)
K.W. McElhane, J.J. Vlassak, W.D. Nix, *J. Mater. Res.* **13**, 05 (1998)
H.B. Muhlhaus, E.C. Aifantis, *Acta Mech.* **89**, 1–4 (1991)
W.D. Nix, H. Gao, *J. Mech. Phys. Solids* **46**, 3 (1998)
J.F. Nye, *Acta Metall.* **1**, 2 (1953)
W.C. Oliver, G.M. Pharr, *J. Mater. Res.* **7**, 6 (1992)
N.A. Stelmashenko, M.G. Walls, L.M. Brown, Y.V. Milman, *Acta Metall. Mater.* **41**, 10 (1993)
J.S. Stolken, A.G. Evans, *Acta Mater.* **46**, 14 (1998)
D. Tabor, *J. Inst. Met.* **79**, 1 (1951)
I. Vardoulakis, E.C. Aifantis, *Acta Mech.* **87**, 3–4 (1991)
A.A. Volinsky, N.R. Moody, W.W. Gerberich, *J. Mater. Res.* **19**, 9 (2004)
G.Z. Voyiadjis, R.K. Abu Al-Rub, *Int. J. Plast.* **19**, 12 (2003)
G.Z. Voyiadjis, R.K. Abu Al-Rub, *Int. J. Solids Struct.* **42**, 14 (2005)

- G.Z. Voyiadjis, A.H. Almasri, *J. Eng. Mech.* **135**, 3 (2009)
- G.Z. Voyiadjis, D. Faghihi, *Procedia ITUTAM*, 3, pp. 205–227 (2012)
- G.Z. Voyiadjis, R. Peters, *Acta Mech.* **211**, 1–2 (2010)
- G.Z. Voyiadjis, C. Zhang, *Mater. Sci. Eng. A* **621**, 218 (2015)
- G.Z. Voyiadjis, D. Faghihi, C. Zhang, *J. Nanomech Micromech* **1**, 1 (2011)
- Y. Xiang, J.J. Vlassak, *Scr. Mater.* **53**, 2 (2005)
- B. Yang, H. Vehoff, *Acta Mater.* **55**, 3 (2007)
- H.M. Zbib, E.C. Aifantis, *Res. Mechanica.* **23**, 2–3 (1998)
- C. Zhang, G.Z. Voyiadjis, *Mater. Sci. Eng. A* **659**, 55 (2016)



Size Effects During Nanoindentation: Molecular Dynamics Simulation

2

George Z. Voyiadjis and Mohammadreza Yaghoobi

Contents

Introduction	40
Simulation Methodology	41
Boundary Conditions Effects	45
Comparing MD Results with Theoretical Models	50
Size Effects in Small-Length Scales During Nanoindentation	52
Effects of Grain Boundary on the Nanoindentation Response of Thin Films	59
References	74

Abstract

In this chapter, the molecular dynamics (MD) simulation of nanoindentation experiment is revisited. The MD simulation provides valuable insight into the atomistic process occurring during nanoindentation. First, the simulation details and methodology for MD analysis of nanoindentation are presented. The effects of boundary conditions on the nanoindentation response are studied in more detail. The dislocation evolution patterns are then studied using the information provided by atomistic simulation. Different characteristics of metallic sample during nanoindentation experiment, which have been predicted by theoretical models, are investigated. Next, the nature of size effects in samples with small length scales are studied during nanoindentation. The results indicate that the size effects at small indentation depths cannot be modeled using the forest

G. Z. Voyiadjis (✉) · M. Yaghoobi
Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge,
LA, USA
e-mail: voyiadjis@eng.lsu.edu; myaghol@lsu.edu

hardening model, and the source exhaustion mechanism controls the size effects at the initial stages of nanoindentation. The total dislocation length increases by increasing the dislocation density which reduces the material strength according to the exhaustion hardening mechanisms. The dislocation interactions with each other become important as the dislocation content increases. Finally, the effects of grain boundary (GB) on the controlling mechanisms of size effects are studied using molecular dynamics.

Keywords

Nanoindentation · Molecular dynamics · Size effects · Dislocation · Grain boundary

Introduction

Indentation is a common experiment to investigate the material properties at different length scales. During indentation, a required force to press a hard indenter into the sample is measured. In the case of nanoindentation, unlike the traditional indentation experiment, it has been observed that the hardness is not a constant value and varies during the test (Nix and Gao 1998; Al-Rub and Voyiadjis 2004; Voyiadjis and Al-Rub 2005). Many researchers have tried to study the variation of hardness, which is commonly termed as size effects, during nanoindentation. The variation of geometrically necessary dislocations (GNDs) density has been commonly considered as the mechanism which controls the hardness. Corcoran et al. (1997) investigated the dislocation nucleation and its effects on the response of Au during nanoindentation experiment. Suresh et al. (1999) studied the effects of sample thickness on the mechanical and dislocation nucleation of Cu thin films during nanoindentation. The grain boundary (GB) effects on the defect nucleation and evolution of bicrystal Fe-14 wt. %Si alloy during nanoindentation were investigated by Soer and De Hosson (2005). It was observed that the dislocations pile up against the GB (Soer and De Hosson 2005). Almasri and Voyiadjis (2010) conducted the nanoindentation of polycrystalline thin films and observed that the GB may enhance the sample hardness.

The interaction of dislocations with each other governs the material strength in bulk metallic samples which are usually captured by Taylor-like hardening models (Nix and Gao 1998; Al-Rub and Voyiadjis 2004; Voyiadjis and Al-Rub 2005). The models generally relate the strength to the dislocation density and state that the stress increases by increasing the dislocation density. Recently, researchers have been able to experimentally measure the GNDs content in samples of confined volume (Kysar and Briant 2002; Kysar et al. 2007; Zaafarani et al. 2008; Demir et al. 2009; Dahlberg et al. 2014). However, the experimental observations cannot be fully described by the bulk size models. Demir et al. (2009, 2010) conducted the nanoindentation and microbending experiments and observed that the governing mechanisms of size effects at smaller length scales

are not similar to those of the large size samples. Demir et al. (2009) observed that the hardness decreases by increasing the GNDs density during nanoindentation of Cu single crystal thin films which cannot be described using the bulk-sized models. Demir et al. (2010) also showed the breakage of the dislocations mean-field theory at small length scales during the microbending of Cu single crystal thin films.

One approach to investigate the governing atomistic process of size effects during nanoindentation is to simulate the sample with the full atomistic details using MD. Many deformation mechanisms during nanoindentation of metallic thin films have been captured using MD. Incorporating the MD simulation, Kelchner et al. (1998) investigated the defect nucleation and evolution of Au during nanoindentation. The surface step effects on the response of Au during nanoindentation were investigated by Zimmerman et al. (2001) using atomistic simulation. Lee et al. (2005) conducted a comprehensive study on the defect nucleation and evolution patterns during nanoindentation of Al and tried to explain the nanoindentation response using those patterns. Hasnaoui et al. (2004), Jang and Farkas (2007), and Kulkarni et al. (2009) have studied the interaction between the dislocations and GB during nanoindentation experiment using molecular dynamics. Yaghoobi and Voyiadjis (2014) investigated the effects of the MD boundary conditions on the sample response and defect nucleation and evolution patterns during nanoindentation. Voyiadjis and Yaghoobi (2015) investigated the theoretical models developed to capture the size effects during nanoindentation using MD. Yaghoobi and Voyiadjis (2016a) investigated the governing mechanisms of size effects during nanoindentation using MD. Voyiadjis and Yaghoobi (2016) incorporated large-scale MD to study the GB effects on the material strength as the grain size varies.

This chapter is designed as follows. In section “Simulation Methodology,” the general details for atomistic simulation of nanoindentation are described. In section “Boundary Conditions Effects,” the effects of selected boundary conditions for MD simulation of nanoindentation are elaborated. In section “Comparing MD Results with Theoretical Models,” the obtained results from MD simulation are compared to those predicted by the available theoretical models. In section “Size Effects in Small-Length Scales During Nanoindentation,” the governing mechanisms of size effects in thin films of confined volumes are presented. In section “Effects of Grain Boundary on the Nanoindentation Response of Thin Films,” the effects of grain boundary and grain size on the nanoindentation response of thin films are elaborated using the results obtained from MD simulation.

Simulation Methodology

The Newton’s equations of motion for N interacting monoatomic molecules can be described as below:

$$m_i \ddot{\mathbf{r}}_i = -\nabla_i U + \mathbf{f}_i, i = 1, 2, \dots, N \quad (1)$$

where $\ddot{\mathbf{r}}_i$ is the second time derivative of i^{th} particle trajectory \mathbf{r}_i , m_i is the mass of i^{th} particle, \mathbf{f}_i is an external force on the i^{th} particle, and

$$\nabla_i U = \left(\frac{\partial}{\partial x_i} + \frac{\partial}{\partial y_i} + \frac{\partial}{\partial z_i} \right) U, i = 1, 2, \dots, N \quad (2)$$

where $U(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ is the potential energy. A metallic system can be described using Eq. 1 by approximating the atoms as mass points. Equation 1 should be numerically solved for the whole system. In the case of atomistic simulation, N can be a large number, and a very efficient parallel code should be used to solve the equation. The atomic interactions in metallic systems have been modeled using many different potentials such as Lennard–Jones (LJ), Morse, embedded-atom method (EAM), and modified embedded-atom method (MEAM).

The LJ potential E^{LJ} can be described as below:

$$E^{\text{LJ}}(r_{ij}) = 4\varepsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right] \quad (3)$$

where σ is the distance from the atom at which $E^{\text{LJ}} = 0$ and ε is the potential well depth. A cutoff distance should be chosen for LJ potential.

Morse potential E^{Morse} can be written as follows:

$$E^{\text{Morse}}(r_{ij}) = D \left\{ e^{-2\alpha(r_{ij}-r_0)} - 2e^{-\alpha(r_{ij}-r_0)} \right\} \quad (4)$$

where D is the cohesive energy, α is the elastic parameter, and r_0 is the equilibrium distance.

EAM is a popular potential to model the metallic atoms interactions (Daw and Baskes 1984). The EAM potential E^{EAM} is described as below:

$$E^{\text{EAM}}(r_{ij}) = \frac{1}{2} \sum_{i,j} V(r_{ij}) + \sum_i F(\rho_i), \quad \rho_i = \sum_{i \neq j} \varphi(r_{ij}) \quad (5)$$

where $V(r_{ij})$ is the pair interaction potential, $F(\rho_i)$ denotes the embedding potentials, and $\varphi(r_{ij})$ is a function which is defined using the electron charge density.

The MEAM potential is the modification of EAM potential which can be described as follows (Baskes 1992):

$$E^{\text{MEAM}}(r_{ij}) = \frac{1}{2} \sum_{i,j} V(r_{ij}) + \sum_i F(\rho'_i) \quad (6)$$

where

$$\rho'_i = \sum_{i \neq j} \varphi(r_{ij}) + \frac{1}{2} \sum_{k,j \neq i} f_{ij}(r_{ij}) f_{ik}(r_{ik}) g_i(\cos \theta_{ijk}) \quad (7)$$

θ_{ijk} is the angle between the i^{th} , j^{th} , and k^{th} atoms. The explicit three-body term is included by introducing the functions f_{ij} , f_{ik} , and g_i .

The indenter can be modeled as a cluster of atoms. However, in order to manage the computational costs, the indenter is commonly modeled as a repulsive potential. Two types of repulsive models can be used to simulate the interaction between the indenter and thin film atoms:

- First, the spherical indenter is modeled using an indenter repulsive force which is described as below (Yaghoobi and Voyiadjis 2014):

$$\begin{aligned} F^{\text{ind}}(r) &= -K^{\text{ind}}(r - R)^2 & \text{for } r < R \\ F^{\text{ind}}(r) &= 0 & \text{for } r \geq R \end{aligned} \quad (8)$$

where K^{ind} is the force constant, r is the atomic distance to the indenter surface, and R is the radius of indenter.

- Second, the indenter, with general geometry can be simplified using a repulsive potential which is described as below (Voyiadjis and Yaghoobi 2015):

$$E^{\text{ind}}(r) = \varepsilon^{\text{ind}}(r - r_c)^2 r < r_c \quad (9)$$

where ε^{ind} is the force constant, r is the distance between the particle and indenter surface, and r_c is the cutoff distance.

Five different indenter geometries of right square prismatic, spherical, cylindrical, blunt conical, and conical with the spherical tip are incorporated in this chapter. During nanoindentation, the precise contact area (A) should be captured to calculate the hardness at each step. A 2D-mesh is produced from the projections of atoms in contact with the indenter. The total contact area is then calculated using the obtained 2-D mesh.

The true indentation depth h is different from the tip displacement d during nanoindentation. A conical indenter of h can be obtained as below:

$$h = \frac{(a_c - a_0)}{\tan \theta} \quad (10)$$

where θ is the cone semi-angle, $a_c = \sqrt{A/\pi}$ is the contact radius, and $a_0 = r_2 + r_c(1/\cos \theta - \tan \theta)$. The indentation depth of a spherical indenter is obtained as below:

$$h = R - \sqrt{R^2 - a_c^2} \quad (11)$$

In the case of a conical indenter with the spherical tip, h can be calculated using Eq. 11 for the spherical part. The indentation depth can be obtained as follows for the conical part:

$$h = \frac{(a_c - a_0)}{\tan(\theta/2)} + h_0 \quad (12)$$

where h_0 is the depth at which the indenter geometry changes from spherical to conical, and a_0 is the contact radius at h_0 (Fig. 1). In the cases of cylindrical and right square prismatic indenters, it is assumed that $h \approx d$ because there is no relation between indentation depth and contact area.

In order to visualize the defects, several methods have been introduced such as energy filtering, bond order, centrosymmetry parameter, adaptive common neighbor analysis, Voronoi analysis, and neighbor distance analysis which have been compared with each other by Stukowski (2012). Also, the Crystal Analysis Tool developed by Stukowski and his coworkers (Stukowski and Albe 2010; Stukowski et al. 2012; Stukowski 2012, 2014) have been incorporated to extract the dislocations from the atomistic data. Here, centrosymmetry parameter (CSP) and Crystal Analysis Tool are explained in more detail. CSP can be described as below (Kelchner et al. 1998):

$$CSP = \sum_{i=1}^{N_p} |\mathbf{R}_i + \mathbf{R}_{i+N_p}|^2 \quad (13)$$

where \mathbf{R}_i and \mathbf{R}_{i+N_p} are vectors from the considered atom to the i^{th} pair of neighbors, and N_p depends on the crystal structure. For example, $N_p = 6$ for fcc

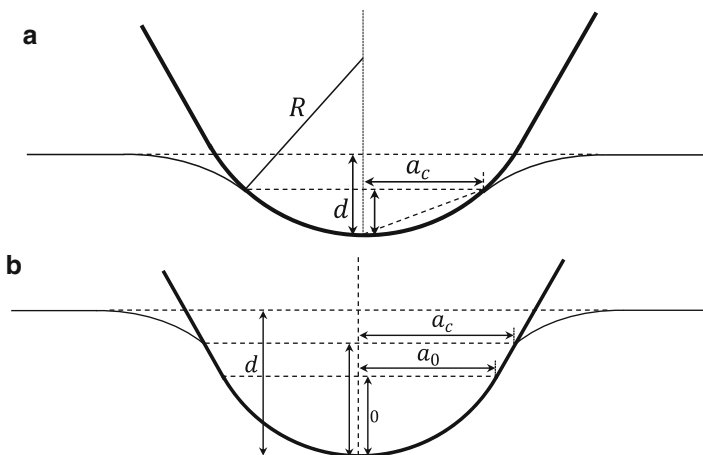


Fig. 1 The true indentation depth h : **a** spherical part of the indenter **b** conical part of the indenter (Reprinted with permission from Yaghoobi and Voyiadjis 2016a)

materials. CSP is equal to zero for perfect crystal structure. However, the atomic vibration introduces a small CSP for atoms which are not defects. Accordingly, a cutoff should be introduced in a way that if $CSP_i < CSP_{\text{cutoff}}$, the i^{th} atom is not considered as a defect (Yaghoobi and Voyiadjis 2014). Also, point defect is removed to clearly illustrate stacking faults. Second, the MD outputs can be postprocessed using the Crystal Analysis Tool (Stukowski and Albe 2010; Stukowski 2012, 2014; Stukowski et al. 2012). The common-neighbor analysis method (Faken and Jonsson 1994) is the basic idea of this code. The code is able to calculate the dislocation information such as the Burgers vector and total dislocation length. To extract the required information, the Crystal Analysis Tool constructs a Delaunay mesh which connects all atoms. Next, using the constructed mesh, the elastic deformation gradient tensor is obtained. The code defines the dislocations using the fact that the elastic deformation gradient does not have a unique value when a tessellation element intersects a dislocation.

Boundary Conditions Effects

One of the most important parts of the MD simulation is to select the appropriate boundary conditions which can accurately mimic the considered phenomenon. In the case of nanoindentation, the selected boundary conditions may influence the response of the simulated material. Up to now, four different boundary conditions types have been incorporated in MD to simulate the nanoindentation experiment which can be described as below (Fig. 2):

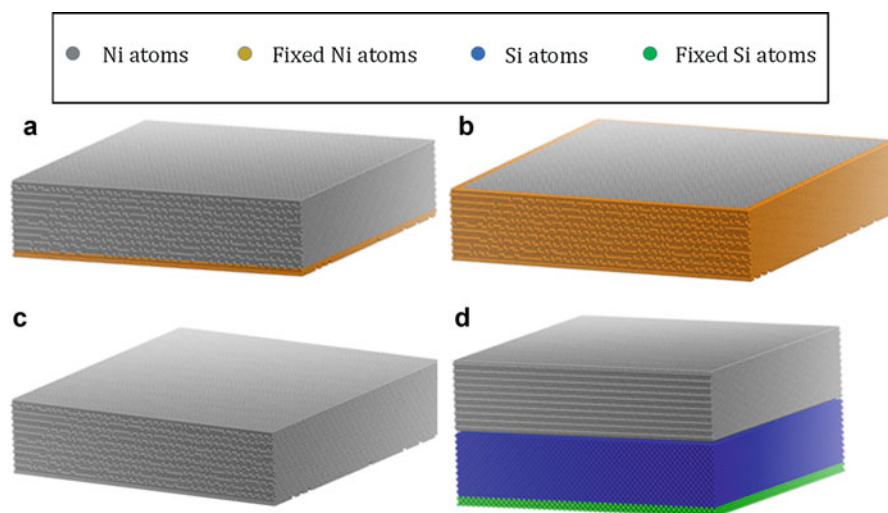


Fig. 2 Boundary conditions of thin films **a** BC1, **b** BC2, **c** BC3, and **d** BC4 (Reprinted with permission from Yaghoobi and Voyiadjis 2014)

- *BC1*: Fixing some atomic layers at the sample bottom to act as a substrate, using free surface for the top, and periodic boundary conditions for the remaining surfaces (see e.g., Nair et al. 2008; Kelchner et al. 1998; Zimmerman et al. 2001)
- *BC2*: Fixing some atomic layer at the surrounding surfaces and using free surfaces for the sample top and bottom (see e.g., Medyanik and Shao 2009; Shao and Medyanik 2010)
- *BC3*: Using free surface for the sample top and bottom, incorporating the periodic boundary conditions for the remaining surfaces, and putting a substrate under the thin film (see e.g., Peng et al. 2010)
- *BC4*: Incorporating the free surfaces for the sample top and bottom, using periodic boundary conditions for the remaining surfaces, and equilibrating the sample by adding some forces (see e.g., Li et al. 2002; Lee et al. 2005)

Yaghoobi and Voyiadjis (2014) studied the different types of boundary conditions and their effects on the dislocation nucleation and evolution patterns using samples with various thicknesses (t_f) indented by spherical indenters with different radii (R). The parallel code LAMMPS (Plimpton 1995), which was developed at Sandia National Laboratories, was selected to conduct the MD simulation. The numerical time integration of Eq. 1 was performed using the velocity Verlet algorithm. Three different types of interatomic interaction were incorporated:

- The interaction of Nickel atoms with each other (Ni-Ni)
- The interaction of Silicon atoms with each other (Si-Si)
- The interaction of Nickel atoms with Silicon ones (Ni-Si)

The three interactions were modeled using the embedded-atom method (EAM) potential for Ni-Ni interaction, Tersoff potential for Si-Si interaction, and Lennard–Jones (LJ) potential for Ni-Si interaction. The Ni-Ni interaction is modeled using the EAM potential parameterized by Mishin et al. (1999). To capture the Si-Si interaction, a three-body Tersoff potential (Tersoff 1988) was chosen. Table 1 presents the potential parameters of Si. The LJ potential E^{LJ} was used to model the Ni-Si interaction and required parameters ($\epsilon_{\text{Ni-Si}}$ and $\sigma_{\text{Ni-Si}}$) are presented in Table 2. The cutoff distance of 2.5σ was selected for LJ potential. The indenter was spherical modeled using the repulsive force F^{ind} presented in Eq. 8. The centrosymmetry parameter (*CSP*) was incorporated to visualize the defects with the cutoff equal to 1.5 ($CSP_{\text{cutoff}} = 1.5$).

Table 1 Tersoff potential parameters of Si-Si (Yaghoobi and Voyiadjis 2014)

$A = 3264.7 \text{ eV}$	$B = 95.373 \text{ eV}$	
$\lambda_1 = 3.2394 \text{ \AA}^{-1}$	$\lambda_2 = 1.3258 \text{ \AA}^{-1}$	
$\alpha = 0$	$\beta = 0.33675$	$n = 22.956$
$c = 4.8381$	$d = 2.0417$	$h = 0.0000$
$\lambda_3 = \lambda_2$	$R = 3.0 \text{ \AA}$	$D = 0.2 \text{ \AA}$

Table 2 LJ potential parameters of Ni-Ni, Si-Si, and Ni-Si (Yaghoobi and Voyiadjis 2014)

	ϵ (J)	σ (Å)
Ni-Ni	8.3134e-20	2.2808
Si-Si	2.7904e-21	3.8260
Ni-Si	1.5231e-20	3.0534

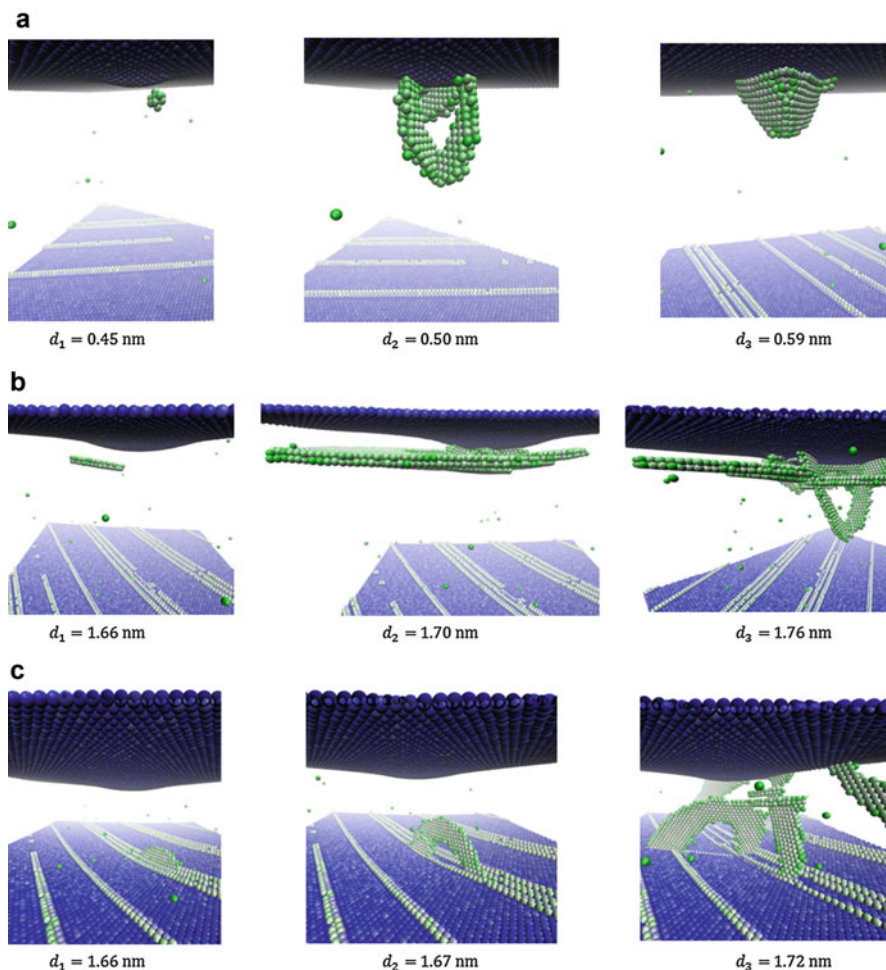


Fig. 3 Defect nucleation and evolution of **a** Type I, **b** Type II, and **c** Type III (Reprinted with permission from Yaghoobi and Voyiadjis 2014)

It was observed that the bending and indentation mechanisms control the initial stages of defect nucleation and evolution for samples with different thicknesses (Yaghoobi and Voyiadjis 2014). Accordingly, three patterns of so-called *Type I*, *Type II*, and *Type III* were observed which can be described as follows (Fig. 3) (Yaghoobi and Voyiadjis 2014):

- *Type I*: The location of initial defect nucleation is beneath the indenter. Two faces of embryonic dislocation loops evolve on $(\bar{1}\bar{1}1)$ and $(1\bar{1}\bar{1})$ planes. Eventually, a tetrahedral sessile lock is formed. This dislocation pattern is controlled by the indentation mechanism.
- *Type II*: Again, the location of initial nucleation is beneath the indenter. However, the defect evolution occurs on the plane which is parallel to (111) , which is the indentation plane. A dislocation pattern similar to *Type I* starts evolving as the indentation depth increases. Both bending and indentation mechanisms are important in this pattern.
- *Type III*: The initial dislocation is nucleated at the sample bottom. The dislocations are evolved on $\{111\}$ planes while they are moving towards the sample top. Bending is the dominant mechanism of deformation.

Figure 4 presents the nanoindentation responses of samples visualized in Fig. 3. In the case of sample with *Type I* defect structure pattern, the first load relaxation occurs due to the initial defect nucleation beneath the indenter (Fig. 4a). Two faces of embryonic dislocation loops evolve on $(\bar{1}\bar{1}1)$ and $(1\bar{1}\bar{1})$ planes. Another load relation occurred when the tetrahedral sessile lock is shaped. Figure 4b illustrates the effects of defect evolution on the response of the sample with dislocation pattern of *Type II* during nanoindentation. The indentation force is initially relaxed due to the first defect nucleation beneath the indenter. The defect evolution occurs on the plane which is parallel to (111) , which is the indentation plane. Due to the complexity of dislocation evolution pattern and activation of both bending and indentation mechanisms, the effects of defects pattern on the nanoindentation response become complicated. Figure 4c shows that indentation load is initially released due to the first dislocation nucleation at the bottom for sample with *Type III* pattern. After the initial load relaxation, the complex pattern of dislocation nucleation and evolution leads to the oscillatory response. However, the general trend is the indentation load increases during indentation.

The governing mechanisms of deformation also depend on the film thickness t_f and indenter radius R (Yaghoobi and Voyiadjis 2014). Samples with BC1 experience no bending independent of the film thickness or indenter radius which leads to the *Type I* pattern controlled by indentation mechanism. Sample with BC2 and BC3 may experience all the patterns of *Type I*, *Type II*, and *Type III* depending on the value of R/t_f . For very small values of R/t_f , the indentation governs the deformation mechanism and *Type I* pattern occurs. Increasing R/t_f , bending mechanism also becomes important and *Type II* pattern occurs. Further increasing R/t_f leads to the *Type III* pattern which is governed by the bending mechanism.

The contact pressure at the onset of plasticity p_m^y is one of the properties which is commonly investigated during nanoindentation. Yaghoobi and Voyiadjis (2014) showed that the p_m^y is also influenced by the choice of MD boundary conditions. The effect can be predicted using the pattern of dislocation nucleation and evolution. In the case of *Type I* pattern, p_m^y is independent of film thickness. However, p_m^y depends on the film thickness for samples with *Type II* and *Type III* in a way that p_m^y increases by increasing the film thickness. Comparing the dislocation nucleation and evolution

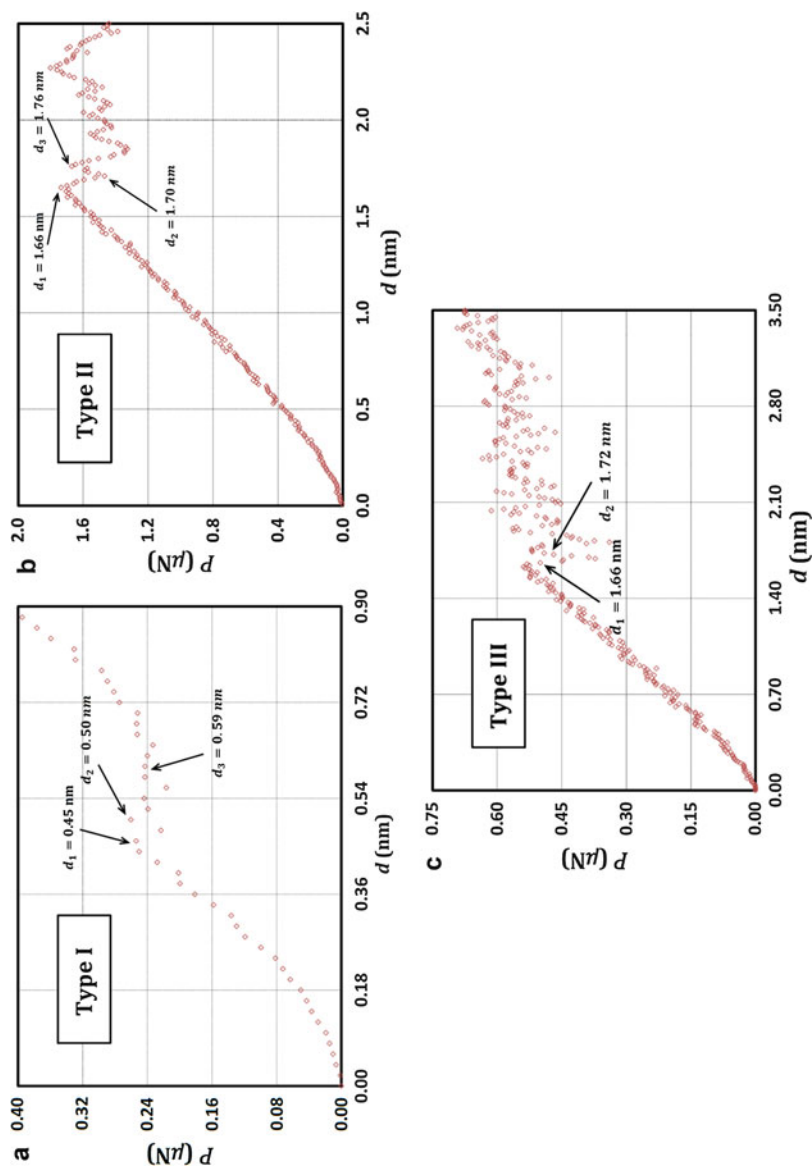


Fig. 4 Nanoindentation responses of thin films with different dislocation nucleation patterns of **a** Type I, **b** Type II, and **c** Type III (Reprinted with permission from Yaghoobi and Voyiadjis 2014)

pattern with each other, *Type I* has the largest and *Type III* has the smallest contact pressure at the time of plasticity initiation.

Comparing MD Results with Theoretical Models

Several theoretical models have been proposed to predict the variation of dislocation length during nanoindentation experiment. Nix and Gao (1998) predicted the variation of geometrically necessary dislocations (GNDs) by integrating the dislocation loops induced during nanoindentation. The dislocation length produced during nanoindentation using a conical indenter can be described as follows (Nix and Gao 1998; Swadener et al. 2002):

$$\lambda_{co} = \frac{\pi a_c h}{b} \quad (14)$$

Pugno (2007) generalized the method by replacing the surface of indentation with the staircase-like surface and presented a generalized equation for dislocation length prediction:

$$\lambda = \frac{S}{b} \quad (15)$$

where $S = \Omega - A$ and Ω is the total contact surface. The variation of dislocation length during nanoindentation using cylindrical and right square prismatic indenters can be obtained using Eq. 15 as below:

$$\lambda_{cy} = \frac{2\pi a_c h}{b} \quad (16)$$

$$\lambda_{pr} = \frac{4ch}{b} \quad (17)$$

where b is the magnitude of the Burgers vector and $c = \sqrt{A}$.

Voyiadjis and Yaghoobi (2015) conducted MD simulation of Ni thin film during nanoindentation to investigate the proposed theoretical models for dislocation length using different indenter geometries of right square prismatic, conical, and cylindrical. The Ni thin film dimensions were 1,200 nm, 1,200 nm, and 600 nm along $[1\bar{1}0]$, $[11\bar{2}]$, and $[111]$ directions, respectively. The radius of cylindrical indenter was $r_1 = 4.8$ nm. The indentation surface of the right square prismatic indenter was a 7.5×7.5 nm². The smaller radius of blunt conical indenter was $r_2 = 0.3$ nm with the cone semi-angle of $\theta = 56.31^\circ$. The parallel code LAMMPS (Plimpton 195) was selected to conduct the MD simulation. The numerical time integration was performed using the velocity Verlet algorithm. *BC4* was incorporated for MD simulation of nanoindentation. The EAM potential is used for Ni-Ni interaction which was parameterized by Mishin et al. (1999). The indenter was modeled

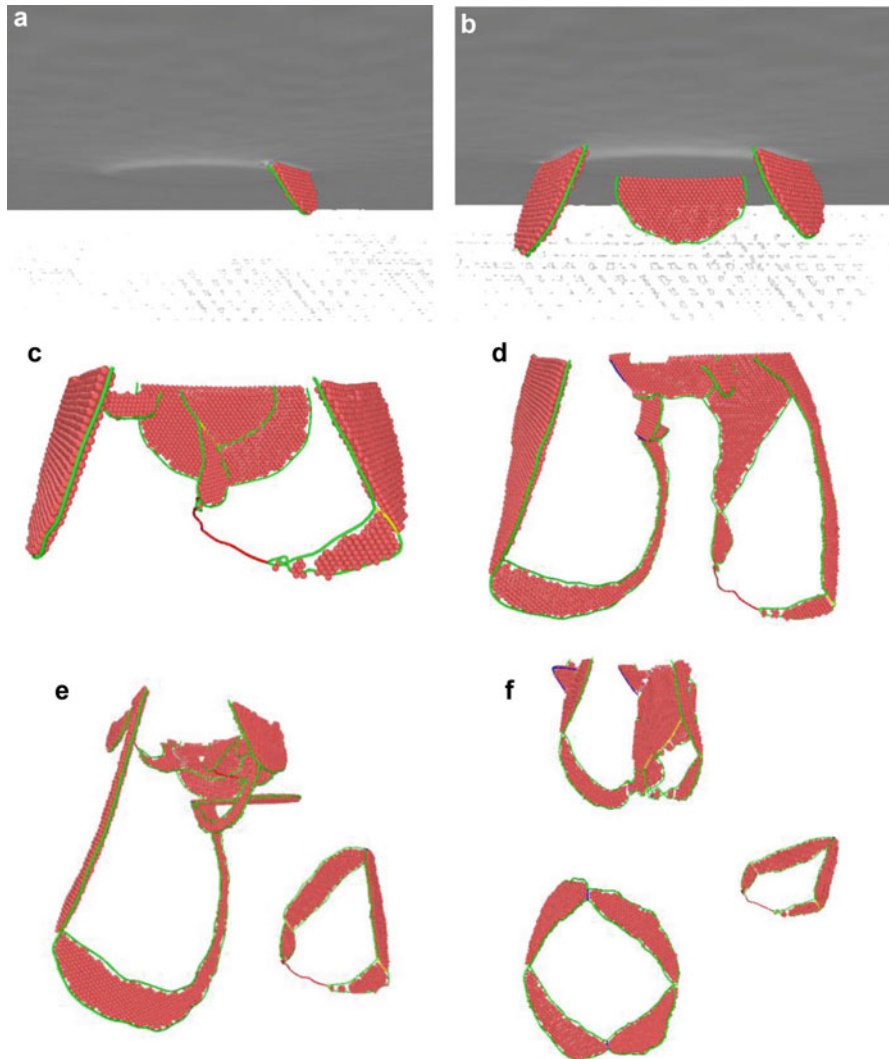


Fig. 5 Defect nucleation and evolution of Ni thin film indented by the cylindrical indenter at **a** $d \approx 0.70$ nm **b** $d \approx 0.86$ nm **c** $d \approx 0.96$ nm **d** $d \approx 1.02$ nm **e** $d \approx 1.05$ nm **f** $d \approx 1.12$ nm (Reprinted with permission from Voyiadjis and Yaghoobi 2015)

using the repulsive potential E^{ind} presented in Eq. 9. The dislocation was extracted from atomistic data using the Crystal Analysis Tool (Stukowski and Albe 2010; Stukowski 2012, 2014; Stukowski et al. 2012).

First, Voyiadjis and Yaghoobi (2015) investigated the dislocation nucleation and evolution pattern during nanoindentation. As an example, Fig. 5 shows the dislocation nucleation and evolution for Ni thin film indented by a cylindrical indenter during nanoindentation. The dislocations and stacking faults are visualized

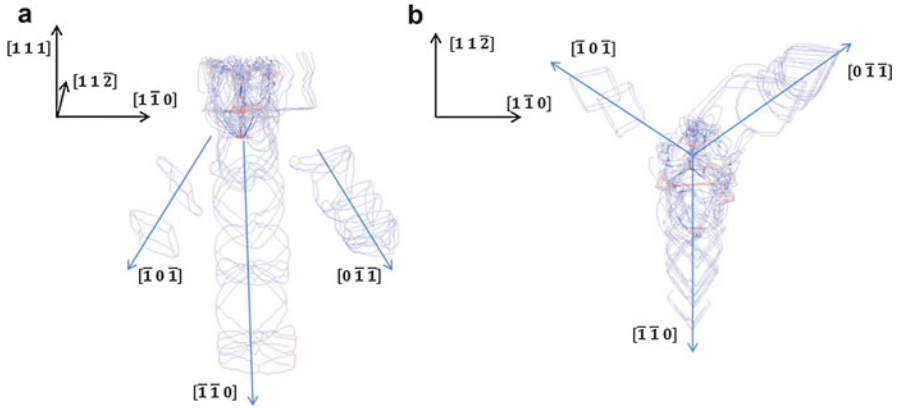


Fig. 6 Prismatic loops forming and movement in Ni thin film indented by the cylindrical indenter during nanoindentation **a** side view **b** top view (Reprinted with permission from Voyiadjis and Yaghoobi 2015)

while the perfect atoms are removed. The color of Shockley, Hirth, and stair-rod partial dislocations and perfect dislocations are green, yellow, blue, and red, respectively. Figure 6 illustrates the dislocation loop formation and movement along three directions of $[\bar{1}0\bar{1}]$, $[\bar{1}\bar{1}0]$, and $[0\bar{1}\bar{1}]$. Figure 7 compares the dislocation lengths obtained from atomistic simulation with those predicted by Eqs. 14–17. The results show that the theoretical predictions can accurately capture the dislocation lengths during nanoindentation. However, some discrepancies are observed which can be described as follows:

- Atomistic simulation captures the total dislocation length including both geometrically necessary and statistically stored dislocations, while the theoretical models only calculate geometrically necessary dislocations. This is the reason that the MD simulation dislocation lengths are mostly higher than that of the theoretical ones.
- The theoretical models incorporate the Burgers vector of the Shockley partial dislocations which comprised most of the dislocation content. However, a few stair-rod and Hirth partial and perfect dislocations are nucleated with the Burgers vectors different from the one for Shockley partial dislocations.
- The dislocations which are detached from the main dislocations network as the prismatic loops and leave the plastic zone around the indenter are not considered in the total dislocation length calculation.

Size Effects in Small-Length Scales During Nanoindentation

Demir et al. (2009) observed that the governing mechanisms of size effects at smaller length scales are not similar to those of the large size samples during nanoindentation. They observed that increasing the dislocation density decreases

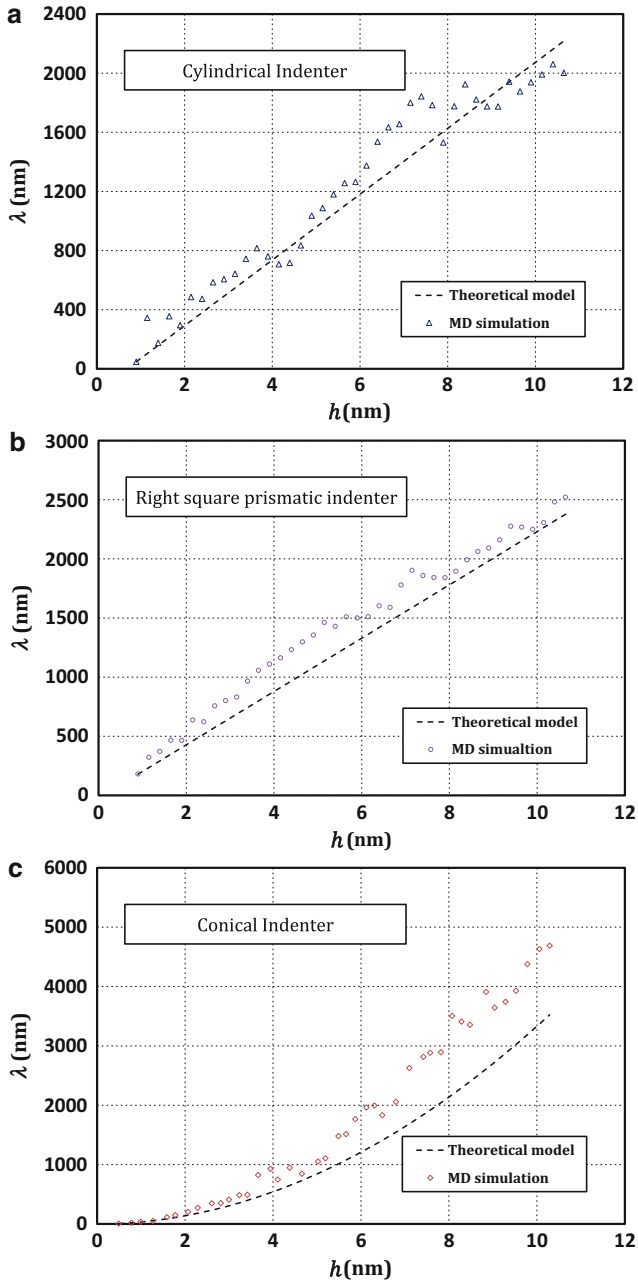


Fig. 7 Total dislocation length obtained from simulation and theoretical models in samples indented by the **a** cylindrical indenter **b** right square prismatic indenter **c** conical indenter (Reprinted with permission from Voyiadjis and Yaghoobi 2015)

the strength (Demir et al. 2009). Uchic et al. (2009), Kraft et al. (2010), and Greer (2013) reviewed the different sources of size effects occurring at smaller length scales. Three mechanisms of source exhaustion, source truncation, and weakest link theory have been introduced to capture the size effects. In samples with very small length scales, lacking enough dislocations to sustain the imposed plastic flow leads to the strength enhancement which is commonly termed as source exhaustion hardening (Rao et al. 2008; El-Awady 2015). The dislocation content reduction may happen due to dislocation starvation, i.e., when the dislocations escape from the sample free surfaces, mechanical annealing, or dislocation source shut down. The material strength also depends on the length of dislocation sources in a way that decreasing the length of dislocation source increases the strength. The dislocation source length becomes smaller by decreasing the sample size through a procedure so-called source truncation. In this procedure, the double-ended dislocation sources transform to a single-ended ones due to the surface effects which decreases the length of dislocation sources. Accordingly, decreasing the sample size leads to a smaller single-ended dislocation source which enhances the material strength (Parthasarathy et al. 2007; Rao et al. 2007). The weakest link theory states that the material strength increases by decreasing the sample size, because decreasing the sample length scale will increase the strength of the weakest source available in the sample (Norfleet et al. 2008; El-Awady et al. 2009).

Yaghoobi and Voyiadjis (2016a) incorporated the large scale MD to study the sources of size effects at smaller length scales during nanoindentation. They selected single crystal Ni thin films with the dimensions of 120 nm, 120 nm, and 60 nm along $[\bar{1}\bar{1}0]$, $[11\bar{2}]$, and $[111]$ directions, respectively. A conical indenter with a spherical tip was incorporated which is similar to the one used by Demir et al. (2009). The remaining simulation methodology is similar to the section “Comparing MD Results with Theoretical Models.”

Swadener et al. (2002) approximated the spherical indenter geometry with a parabola and presented the following equation to predict the GNDs length during nanoindentation:

$$\lambda_{sp} \approx \frac{2\pi}{3} \frac{a_c^3}{bR} \quad (18)$$

The approximation, however, is only applicable for small indentation depths. Yaghoobi and Voyiadjis (2016a) introduced a theoretical equation to predict the dislocation length of sample indented by a spherical tip using the precise geometry of the indenter. The total dislocation length can be described as below (Yaghoobi and Voyiadjis 2016a):

$$\begin{aligned} \lambda_{sp} &= \int_0^{a_c} \frac{2\pi r}{b} \left(\frac{dh}{dr} \right) dr = \frac{2\pi}{b} \int_0^{a_c} \left(\frac{r^2}{\sqrt{R^2 - r^2}} \right) dr \\ &= \frac{2\pi}{b} \left[\frac{R^2}{2} \sin^{-1} \left(\frac{a_c}{R} \right) - \frac{1}{2} \left(a_c \sqrt{R^2 - a_c^2} \right) \right] \end{aligned} \quad (19)$$

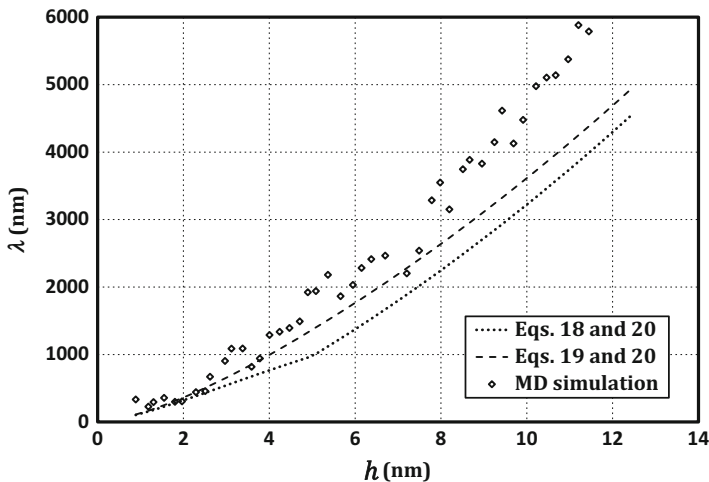


Fig. 8 Comparison between the dislocation lengths obtained from theoretical models and MD simulation during nanoindentation (Reprinted with permission from Yaghoobi and Voyiadjis 2016a)

The GNDs length for the conical part of the indenter can be calculated as below:

$$\lambda_{co} = \lambda_{sp}|_{h=h_0} + \int_{a_0}^{a_c} \frac{2\pi r}{b \tan\left(\frac{\theta}{2}\right)} dr = \lambda_{sp}|_{h=h_0} + \frac{\pi (a_c^2 - a_0^2)}{b \tan\left(\frac{\theta}{2}\right)} \quad (20)$$

Figure 8 compares the dislocation length obtained from atomistic simulation with those calculated from the approximate and precise theoretical models during nanoindentation. The GNDs length calculated from the theoretical model is a lower bound for the total dislocation length obtained from MD which includes all types of dislocations.

A plastic zone should be defined to obtain the dislocation density. Yaghoobi and Voyiadjis (2016a) assumed that the plastic zone is a hemisphere with the radius of $R_{pz} = fa_c$ where f is a constant. The value of $f = 1.9$ was selected by Yaghoobi and Voyiadjis (2016a) for the theoretical dislocation density calculations which is similar to Durst et al. (2005). The density of dislocations can be described as below:

$$\rho = \lambda/V \quad (21)$$

where V is the plastic zone volume. Figure 9 compares the approximate and precise theoretical dislocation densities. Figure 9 shows that approximating a sphere using a parabola leads to a constant dislocation density. However, the density of dislocations increases during nanoindentation by incorporating the precise indenter geometry.

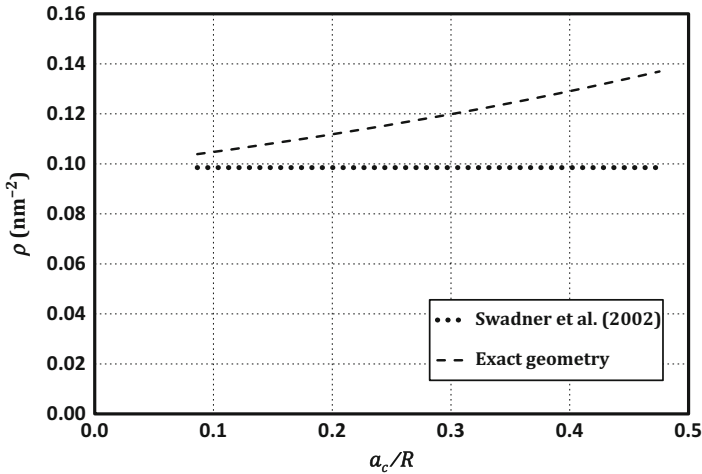


Fig. 9 Theoretical GNDs density obtained from the approximate (Swadner et al. 2002) and exact (see Eq. 19) geometries of the indenter versus the normalized contact radius a_c/R (Reprinted with permission from Yaghoobi and Voyiadjis 2016a)

Yaghoobi and Voyiadjis (2016a) incorporated five different values of $f = 1.5, 2.0, 2.5, 3.0,$ and 3.5 and obtained the corresponding dislocation density for each one. The volume of plastic zone can be obtained by removing the volume occupied by the indenter (V_{indenter}) as follows:

$$V = (2/3) \pi (fa_c)^3 - V_{\text{indenter}} \quad (22)$$

Figure 10 illustrates the variation of dislocation density ρ during nanoindentation for different sizes of plastic zone. The results show that for all values of f , the dislocation density increases during nanoindentation which is in agreement with the trend predicted by the precise theoretical prediction presented in Fig. 9.

Figure 11 presents the variation of the mean contact pressure ($p_m = P/A$), which is equivalent to the hardness H in the plastic region, during nanoindentation. Figure 11 shows that the mean contact pressure follows the Hertzian theory in the elastic region. After the initial dislocation nucleation, however, the results show that the hardness decreases by increasing the indentation depth.

The forest hardening mechanism governs the material strength in bulk-sized samples which relate the material strength to the interaction of dislocations with each other. The famous Taylor hardening-type models are usually incorporated to describe the shear strength in the case of forest hardening mechanism as follows (Voyiadjis and Al-Rub 2005):

$$\tau = \alpha_S \mu b_S \sqrt{\rho}$$

$$\rho = \left[\rho_S^{\beta/2} + (\alpha_G^2 b_G^2 \rho_G / \alpha_S^2 b_S^2)^{\beta/2} \right]^{2/\beta} \quad (23)$$

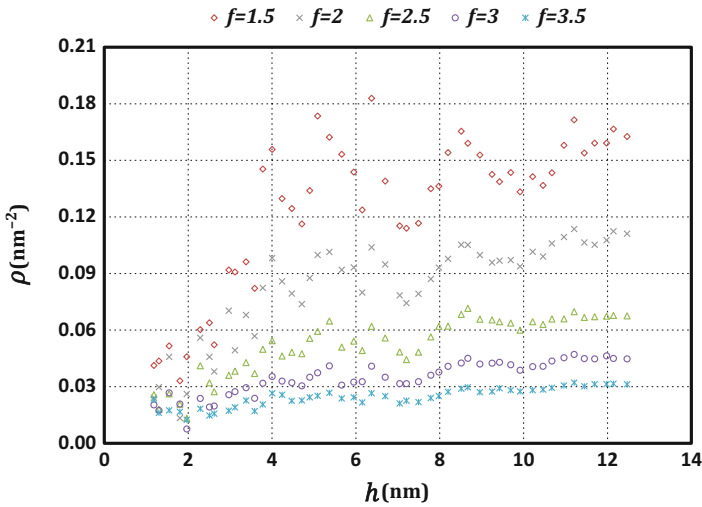


Fig. 10 Dislocation density obtained from MD simulation for different values of f during nanoindentation (Reprinted with permission from Yaghoobi and Voyiadjis 2016a)

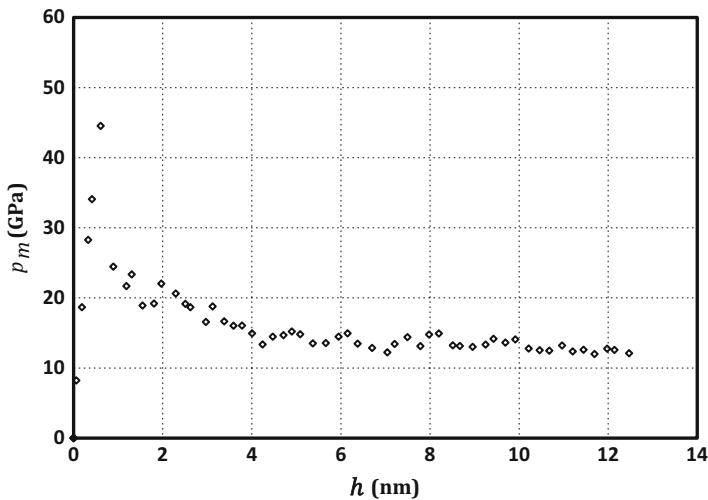


Fig. 11 Variation of mean contact pressure p_m as a function of indentation depth h (Reprinted with permission from Yaghoobi and Voyiadjis 2016a)

where α is a constant, μ is the shear modulus, and the indices G and S designate GNDs and SSDs parameters, respectively. Equation 23 states the material strength increases by increasing the dislocation density. Figure 10 shows that the dislocation density increases by increasing the indentation depth. According to the forest hardening mechanism, the strength should also increase. However, Fig. 11 illustrates

that the hardness decreases during nanoindentation, and the material size effects cannot be captured by the forest hardening mechanism.

The dislocation nucleation and evolution should be investigated in addition to the nanoindentation response to unravel the controlling mechanisms of size effects. Figure 12 depicts the initial stages of dislocation evolution. It shows that the cross-slip is the dominant mechanism to produce the dislocation sources. Elongation of dislocations pinned at their ends provides the required dislocation length to

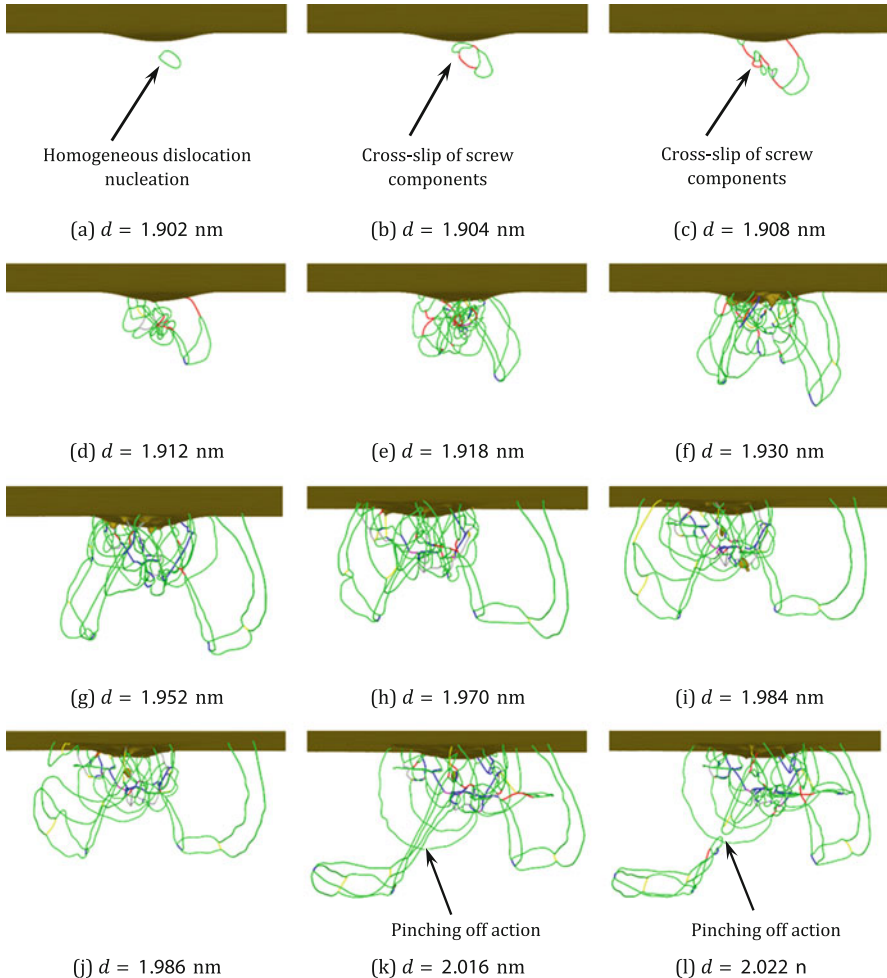


Fig. 12 Dislocation nucleation and evolution at small tip displacements: **a** initial homogeneous dislocation nucleation beneath the indenter which has a Burgers vector of $1/6 [\bar{2}1\bar{1}]$ (Shockley partial dislocation); **b–j** cross-slip of screw components which produces new pinning points; **k–l** first loop is released by pinching off action (Reprinted with permission from Yaghoobi and Voyiadjis 2016a)

sustain the imposed deformation. Dislocation loops are then released by pinching off of screw dislocations and glide along the three directions of $[\bar{1}10]$, $[01\bar{1}]$, and $[\bar{1}0\bar{1}]$. After the initial dislocation nucleation, the available dislocation length is insufficient to sustain the imposed deformation, and the source exhaustion controls the size effects. Consequently, the required stress reduces as the dislocation length and density increase during nanoindentation. The dislocation density and length are eventually reaching the value required to sustain the imposed plastic flow, and the hardness tends to a constant value. Also, the interaction of dislocation with each other becomes important by increasing the dislocation length. However, the dislocation density reaches a constant value and forest hardening mechanism does not lead to any size effects.

Effects of Grain Boundary on the Nanoindentation Response of Thin Films

Grain boundary (GB) has a key role in deformation mechanism of crystalline material (Meyers et al. 2006; Koch et al. 2007; Zhu et al. 2008). For crystalline materials with large grains, the Hall-Petch relation describes the effects of grain size which states the material strength increases by decreasing the grain size. Hall-Petch effect is commonly attributed to the dislocation pile-up mechanism (Meyers et al. 2006; Koch et al. 2007; Zhu et al. 2008). The Hall-Petch relation breaks down for grain size smaller than some limits, and other deformation mechanisms control the size effects in crystalline materials such as the grain boundary rotation and sliding (Meyers et al. 2006; Koch et al. 2007; Zhu et al. 2008).

Atomistic simulation is a powerful tool to study the interaction of dislocations with GBs. Several mechanisms of dislocation reflection, transmission, and absorption were investigated by De Koning et al. (2003) by incorporating the atomistic simulation. Hasnaoui et al. (2004) studied the interaction between the dislocations and GB during nanoindentation experiment using molecular dynamics. Jang and Farkas (2007) conducted the atomistic simulation of bicrystal nickel thin film nanoindentation and observed that the GBs can contribute to the nanoindentation hardness. Kulkarni et al. (2009), however, observed that the GBs mainly reduce the hardness of the metallic samples. They showed that the CTB has the least hardness reduction compared to the other types of GBs (Kulkarni et al. 2009). Tsuru et al. (2010) investigated the effect of indenter distance from the GB using MD. Stukowski et al. (2010) conducted MD simulation of nanoindentation for metallic samples with twin boundaries and observed that the effects of twin GBs on the material response depends on the unstable stacking fault and twin boundary migration energies. Sangid et al. (2011) proposed an inverse relation between the GB energy barrier and GB energy based on the MD simulation results.

Effects of GB on the response of thin film during nanoindentation has been studied by many researchers (Hasnaoui et al. 2004; Jang and Farkas 2007; Kulkarni et al. 2009; Tsuru et al. 2010). However, a study which

addresses a wide range of grain sizes is not a trivial task due to the MD simulation limitations. Voyiadjis and Yaghoobi (2016) incorporated the large-scale MD to study the effects of grain size and grain boundary geometry on the nanoindentation response. They incorporated Ni thin films with two sizes of $24 \times 24 \times 12$ nm (S1) and $120 \times 120 \times 60$ nm (S2). Four symmetric tilt boundaries of $\Sigma 3(111) [1\bar{1}0]$ ($\theta = 109.5^\circ$), $\Sigma 11(113) [1\bar{1}0]$ ($\theta = 50.5^\circ$), $\Sigma 3(112) [1\bar{1}0]$ ($\theta = 70.5^\circ$), and $\Sigma 11(332) [1\bar{1}0]$ ($\theta = 129.5^\circ$) and three asymmetric tilt boundaries of $\Sigma 11(225)/(441)$ ($\varphi = 54.74^\circ$), $\Sigma 3(112)/(\bar{5}52)$ ($\varphi = 19.47^\circ$), and $\Sigma 3(114)/(110)$ ($\varphi = 35.26^\circ$) were generated at the two third of the sample from bottom to compare the governing mechanisms of size effects with those of the single crystal thin films. θ and φ are the interface misorientation and inclination angles, respectively. The spherical indenters with two different radii of $R_1 = 10$ nm and $R_2 = 15$ nm were modeled using the repulsive potential E^{ind} presented in Eq. 9. The procedure to generate and equilibrate the GBs was elaborated by Voyiadjis and Yaghoobi (2016). The equilibrium structures of grain boundaries are illustrated in Fig. 13 using the CSP. The remaining simulation methodology is similar to the section “Comparing MD Results with Theoretical Models.”

Figure 14 depicts the variation of mean contact pressure p_m during nanoindentation for S1 thin films, i.e., the smaller samples. It can be observed that the GB generally reduces the material strength for S1 thin films. However, in the case of coherent twin boundary (CTB), i.e., $\Sigma 3(111) [1\bar{1}0]$, the hardness is slightly enhanced for some indentation depths. Generally, in the cases of smaller thin films, i.e., S1 samples, the CTB has the best performance. Further investigation was conducted by depicting the variation of dislocation length λ during nanoindentation. Figure 15 illustrates the variation of p_m and λ during nanoindentation for CTB. Voyiadjis and Yaghoobi (2016) divided the nanoindentation response to five different regions:

- *Region I:* The bicrystal and single crystal thin films show similar responses during the initial indentation phase which are elastic, and CTB is the only defect that exists in the bicrystal thin film.
- *Region II:* In this region, the dislocation nucleation occurs for the bicrystal thin film beneath the indenter followed by a stress relaxation while the single crystal sample remains elastic. In the case of bicrystal thin film, the size effects is initially governed by the dislocation nucleation and source exhaustion. The dislocation density increases during nanoindentation which decreases the required stress to sustain the imposed plastic flow. Consequently, the hardness decreases by nucleation and evolution of new dislocations.
- *Region III:* The plasticity is initiated in single crystal thin film beneath the indenter followed by a stress relaxation. The thin film strength reduces according to the dislocation nucleation and source exhaustion mechanisms. The dislocation content does not change for bicrystal thin film. Accordingly, the stress should

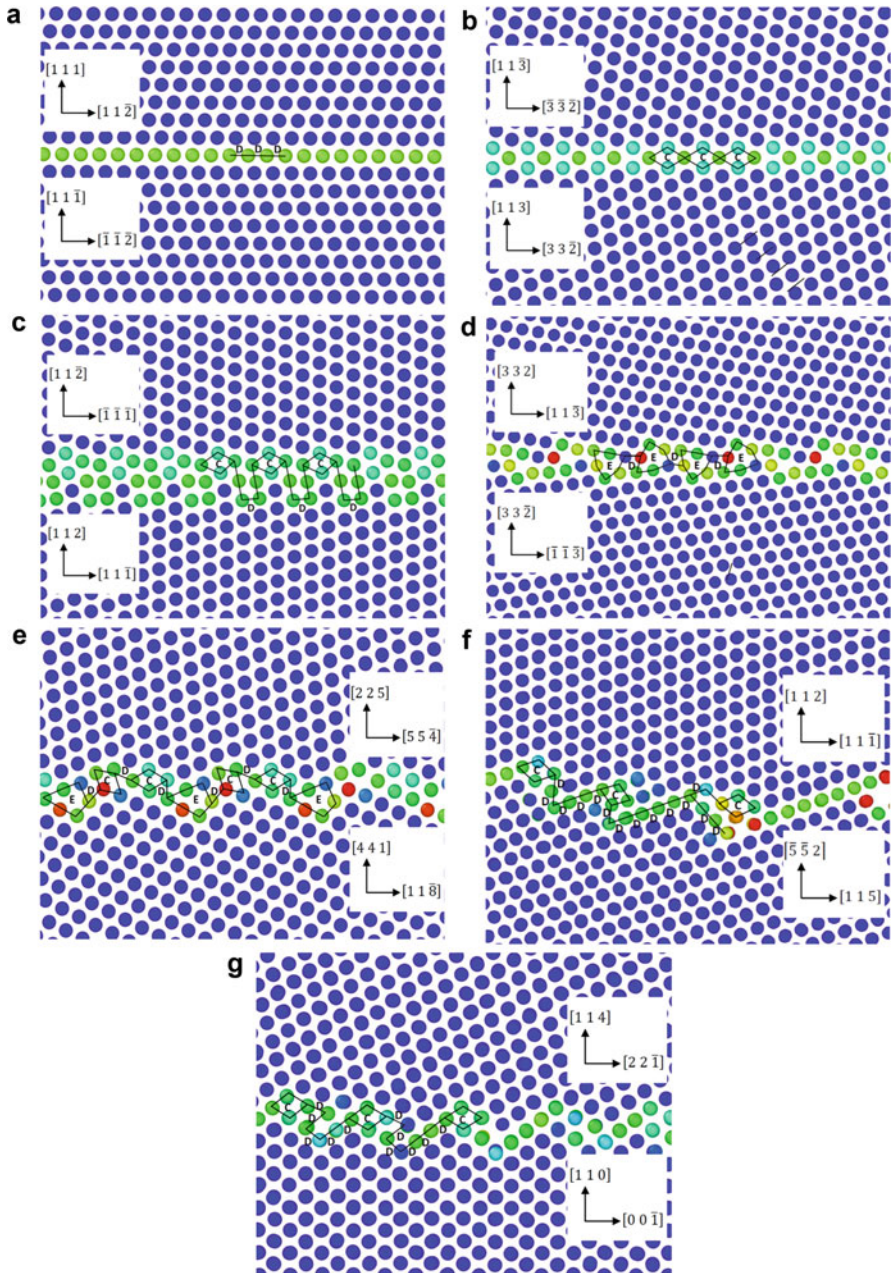


Fig. 13 The equilibrium structure of the symmetric and asymmetric tilt grain boundaries of **a** $\Sigma 3(111) [1\bar{1}0]$ ($\theta = 109.5^\circ$), **b** $\Sigma 11(113) [1\bar{1}0] \Sigma 3(112)$, **c** $[1\bar{1}0]$ ($\theta = 70.5^\circ$), **d** $\Sigma 11(332) [1\bar{1}0]$ ($\theta = 129.5^\circ$), **e** $\Sigma 11(225)/(441)$ ($\varphi = 54.74^\circ$), **f** $\Sigma 3(112)/(\bar{5}52)$ ($\varphi = 19.47^\circ$), and **g** $\Sigma 3(114)/(110)$ ($\varphi = 35.26^\circ$), along $[1\bar{1}0]$ axis (Reprinted with permission from Voyiadjis and Yaghoobi 2016)

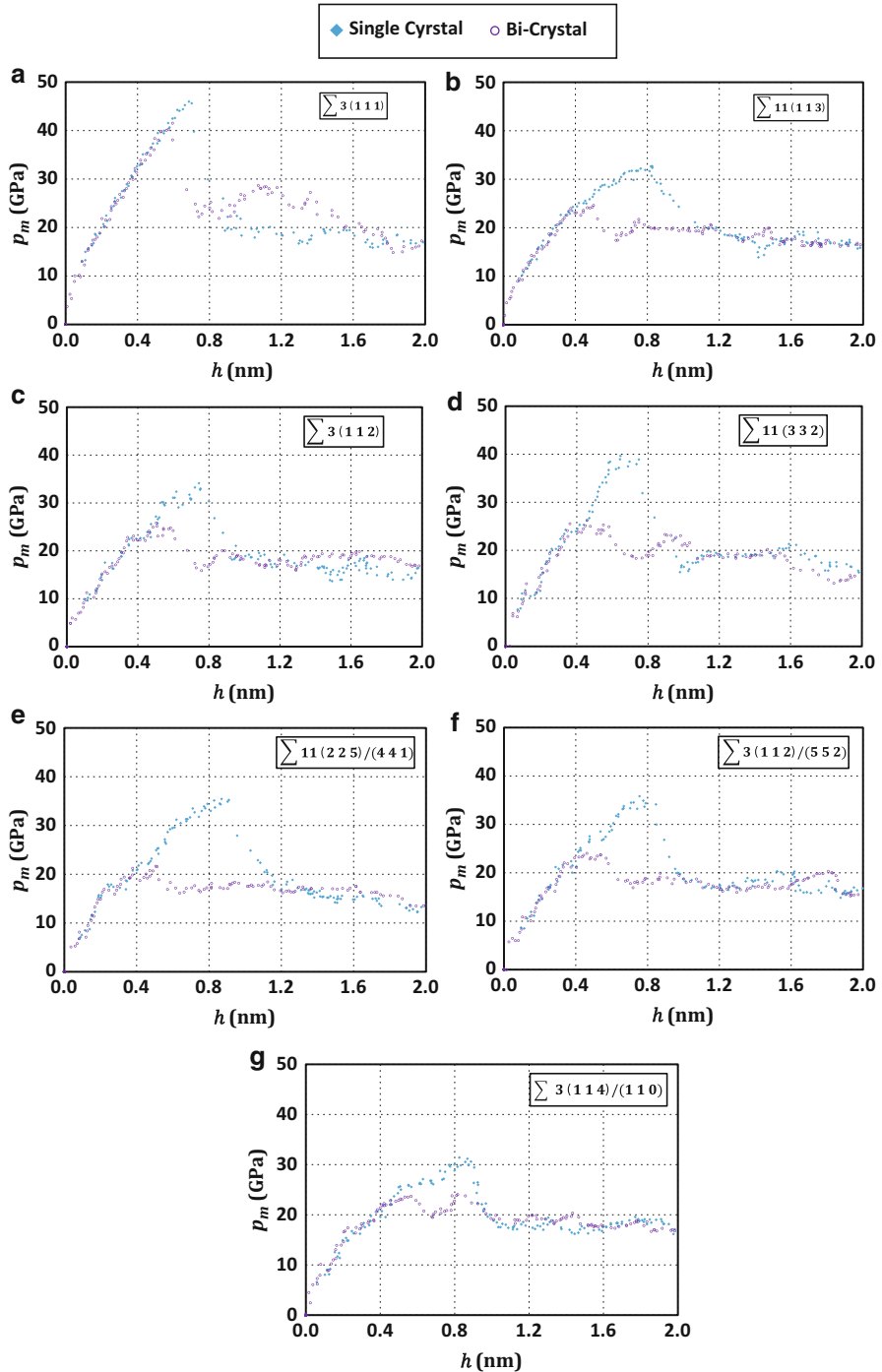


Fig. 14 Variation of mean contact pressure p_m as a function of indentation depth h for S1 single crystal and their related bicrystal samples with the grain boundaries of **a** $\sum 3(111)$, **b** $\sum 11(113)$, **c** $\sum 3(112)$, **d** $\sum 11(332)$, **e** $\sum 11(225)/(441)$, **f** $\sum 3(112)/(552)$, and **g** $\sum 3(114)/(110)$ (Reprinted with permission from Voyiadjis and Yaghoobi 2016)

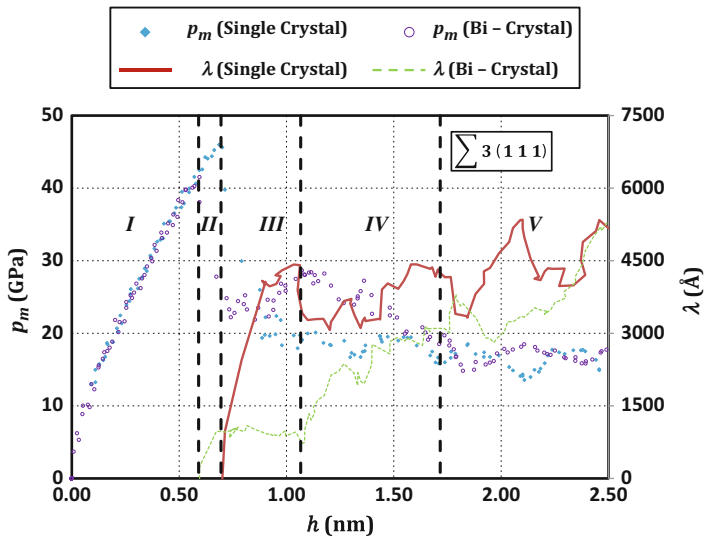


Fig. 15 Variation of mean contact pressure p_m and dislocation length λ as a function of indentation depth h for S1 bicrystal sample with $\Sigma 3(111)$ GB and its related single crystal sample (Reprinted with permission from Voyiadjis and Yaghoobi 2016)

be increased to sustain the imposed deformation based on the source exhaustion mechanism. The dislocations eventually reach the GB which blocks the dislocations. However, the blocked dislocations do not contribute to the strength.

- *Region IV*: The strength in both single and bicrystal thin films is decreased by increasing the dislocation length which follows the source exhaustion mechanism. However, the influence of the source exhaustion mechanism decreases as the dislocation length increases which decreases the slope of the hardness reduction. Also, the dislocations which are blocked by the GB start dissociating into the next grain.
- *Region V*: In this region, the available dislocation content is sufficient to sustain the imposed plastic flow and no further stress reduction occurs. Also, the single and bicrystal thin films reach a similar hardness which shows that the dislocation blockage by GB does not have any contribution to the size effects.

The structures of dislocations in different regions are illustrated in Fig. 16 for bicrystal thin film with CTB and related single crystal sample. Figure 16a, b shows the dislocation structure in *Region II* at which the single crystal sample is defect free and the nucleation occurs beneath the indenter for bicrystal thin film. The results show that the dominant mechanism of dislocation multiplication is cross-slip. Cross-slip introduces the new pinning points and provides the required dislocation length to sustain the plastic flow. Dislocations are elongated while they are pinned at their ends. Figure 16c, d illustrates the dislocation structure in *Region III* while the cross-slip is still the governing mechanism of deformation for both single and

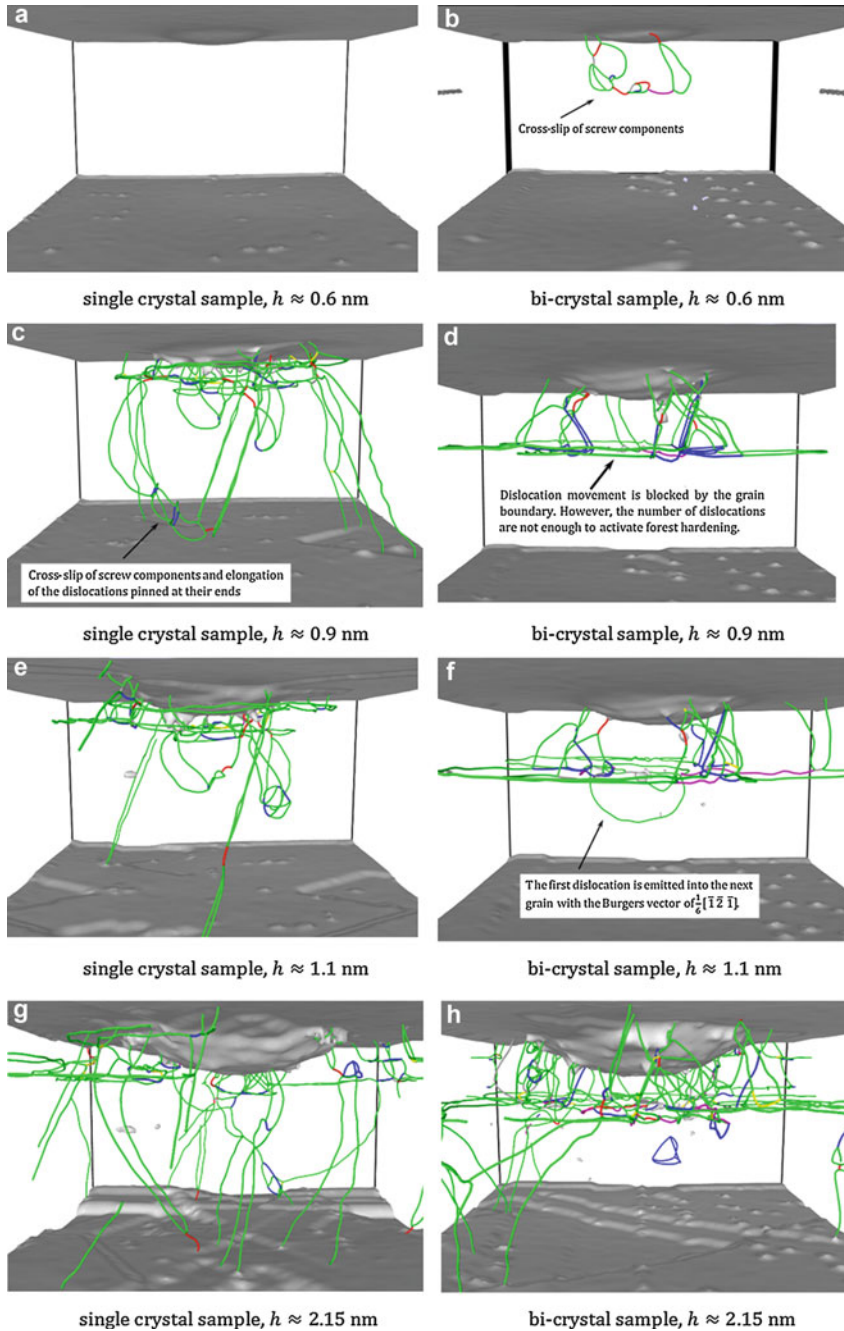


Fig. 16 Dislocation nucleation and evolution: **a** *Region II*, single crystal sample; **b** *Region II*, bicrystal sample; **c** *Region III*, single crystal sample; **d** *Region III*, bicrystal sample; **e** *Region IV*, single crystal sample; **f** *Region IV*, bicrystal sample; **g** *Region V*, single crystal sample; **h** *Region V*, bicrystal sample (Reprinted with permission from Voyiadjis and Yaghoobi 2016)

bicrystal thin films. Figure 16d depicts the dislocation blockage by CTB. Figure 16e, f illustrates the dislocation structure in *Region IV*. Many dislocation multiplications are observed in both single and bicrystal thin films which are induced according to the cross-slip mechanism. Figure 16f shows the initial dislocation dissociation into the next grain in the case of bicrystal sample which is a Shockley partial dislocation with the Burgers vector of $\frac{1}{6} [12\bar{1}]$. In the case of *Region V*, Fig. 16g, h depicts the dislocation structure which shows enough dislocation length is provided to sustain the imposed deformation. Also, the interaction of dislocations with each other cannot be neglected anymore.

Figure 17 shows the variation of p_m and λ during nanoindentation for the S1 samples with different GBs and their related single crystal thin films. In contrast to

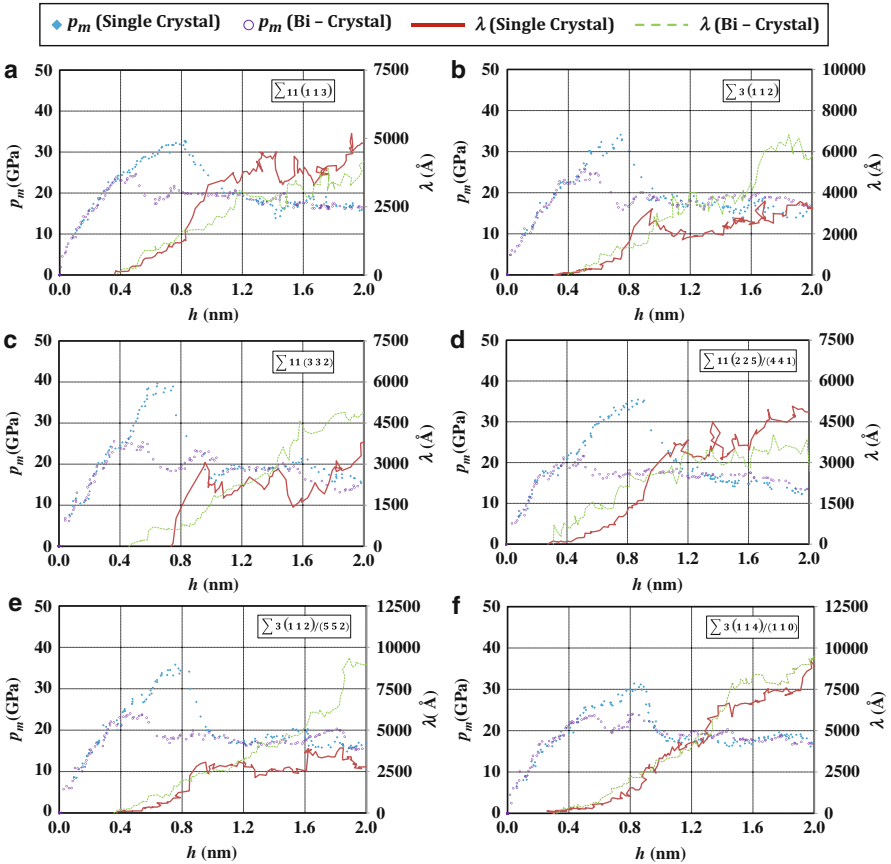


Fig. 17 Variation of mean contact pressure p_m and dislocation length λ as a function of indentation depth h for S1 bicrystal and their related single crystal samples with grain boundaries of: a $\sum_{11}(113)$, b $\sum_3(112)$, c $\sum_{11}(332)$, d $\sum_{11}(225)/(441)$, e $\sum_3(112)/(552)$, and f $\sum_3(114)/(110)$ (Reprinted with permission from Voyiadjis and Yaghoobi 2016)

CTB, Fig. 17 shows that the first stress relaxation does not occur immediately after the first dislocation nucleation. The first large stress relaxation occurs with the first jump in dislocation density for single crystal thin films. The nature of first apparent strength drop in thin films with GB is more complicated due to the interaction of dislocations with GB. Figure 17 shows that the GB decreases the depth at which the first large stress relaxation occurs, and the bicrystal thin films have larger dislocation length at that depth compared to the single crystal samples. The GB itself can be a source of dislocation nucleation which can be activated at different stages of indentation. Figure 18 shows that the GB is the initial source of dislocation nucleation for $\sum 3(112)/(\bar{5}52)$ and $\sum 11(225)/(441)$ GBs, i.e., the initial dislocation nucleation occurs from the GB and not beneath the indenter. The nucleated dislocations are Shockley partial dislocations with the Burgers vectors of $\frac{1}{6} [11\bar{2}]$ and $\frac{1}{6} [112]$ for the GBs of $\sum 3(112)/(\bar{5}52)$ and $\sum 11(225)/(441)$, respectively. If the dislocation nucleation from GB occurs at the initial steps of dislocation nucleation and evolution, it will severely decrease the thin film strength which can be noted for $\sum 3(112)/(\bar{5}52)$ and $\sum 11(225)/(441)$ GBs in Fig. 17d, e. The size effects during nanoindentation can be described for all GBs incorporating the variation of total dislocation length and dislocation visualization during nanoindentation. The results show that the source exhaustion is the controlling mechanism of size effects for the initial stages of dislocation nucleation and evolution. Increasing the total dislocation length, however, the required dislocation length for sustaining the imposed deformation is provided and the source exhaustion mechanism becomes less dominant. Also, the dislocation interactions with each other become nonnegligible by increasing the dislocation content. Eventually, both bicrystal and their related single crystal thin films reach a similar hardness which indicates that the dislocation pileup does not enhance the hardening in the cases of studied S1 samples.

Figure 19 shows mean contact pressure versus indentation depth in the cases of S2 samples, i.e., larger samples. The initial responses of both single crystal and bicrystal thin films are similar. However, GB enhances the hardness for higher indentation depths. In order to unravel the underlying mechanisms of size effects for S2 samples, the variations of mean contact pressure and total dislocation length should be studied. Figure 20 compares the mean contact pressure and total dislocation density of the single crystal thin film with those of the bicrystal sample with CTB. The nanoindentation response can be divided to three different regions:

- *Region I:* There is no plasticity at this region. GB does not change the nanoindentation response of thin film.
- *Region II:* The plasticity is initiated beneath the indenter for both single and bicrystal thin films followed by a sharp stress relaxation. After the initial nucleation, the source exhaustion governs the size effects, and the required stress to maintain the plastic flow decreases by increasing the total dislocation length. In this region, the GB does not significantly change the total dislocation

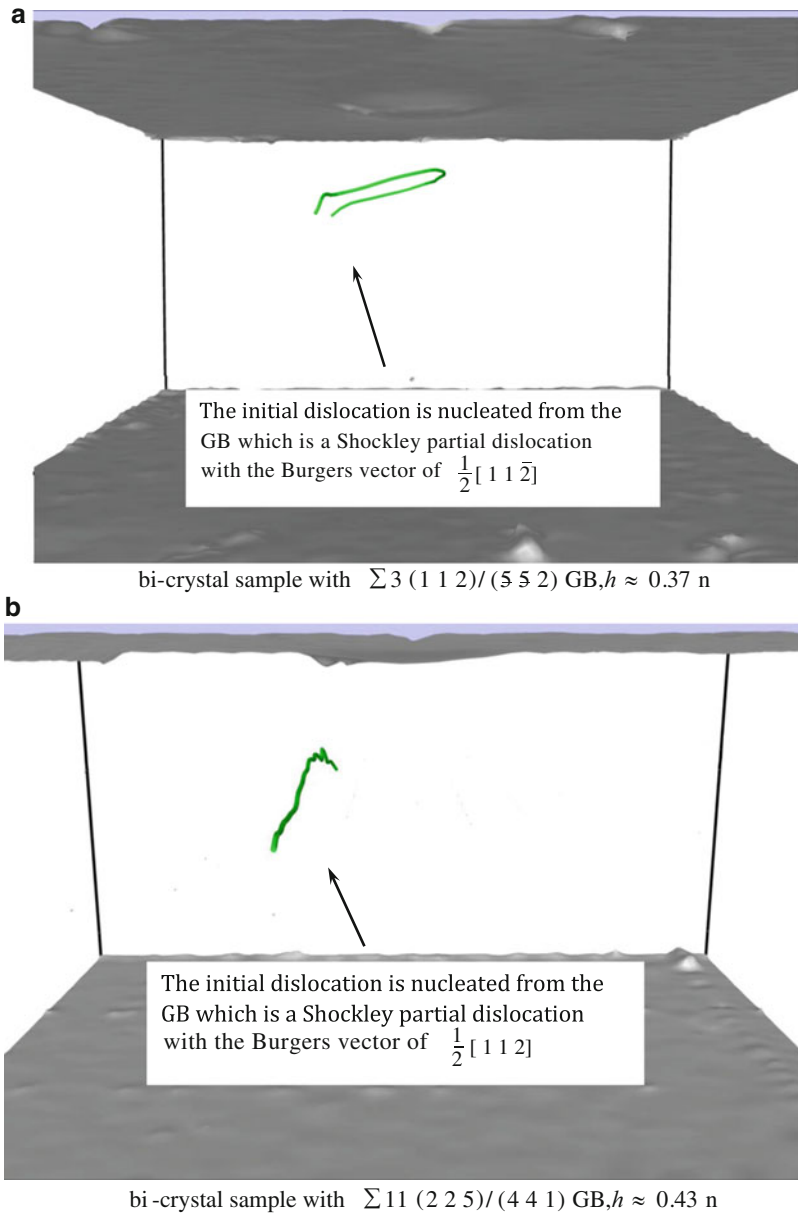


Fig. 18 Dislocation nucleation from the GB: **a** bicrystal sample with $\Sigma 3 (112) / (\bar{5}\bar{5}2)$ GB, $h \approx 0.37$ nm; **b** bicrystal sample with $\Sigma 11(225)/(441)$ GB, $h \approx 0.43$ nm (Reprinted with permission from Voyiadjis and Yaghoobi 2016)

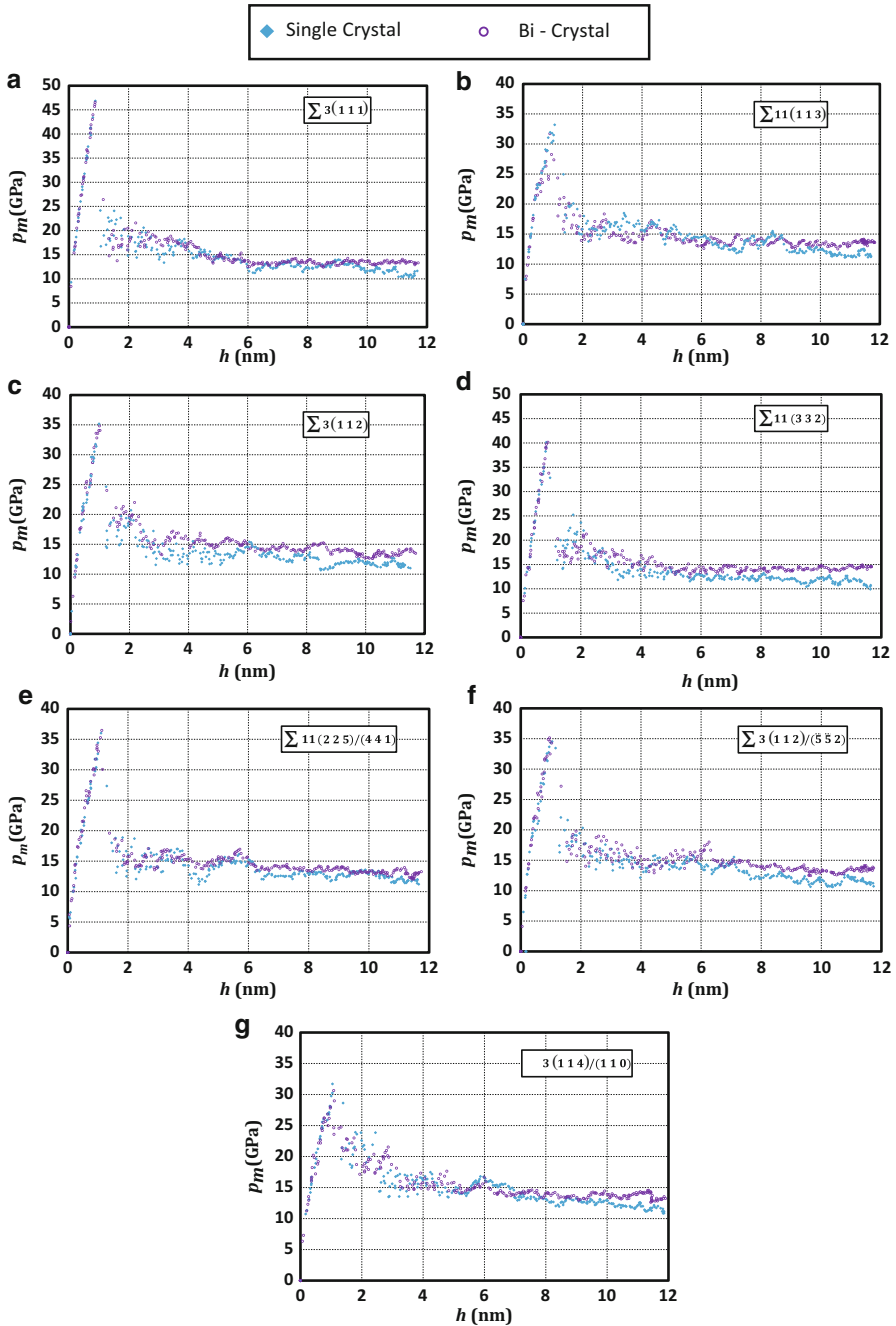


Fig. 19 Variation of mean contact pressure p_m as a function of indentation depth h for S2 single crystal and their related bicrystal samples with the grain boundaries of **a** $\Sigma 3(111)$, **b** $\Sigma 11(113)$, **c** $\Sigma 3(112)$, **d** $\Sigma 11(332)$, **e** $\Sigma 11(225)/(441)$, **f** $\Sigma 3(112)/(552)$, and **g** $\Sigma 3(114)/(110)$ (Reprinted with permission from Voyiadjis and Yaghoobi 2016)

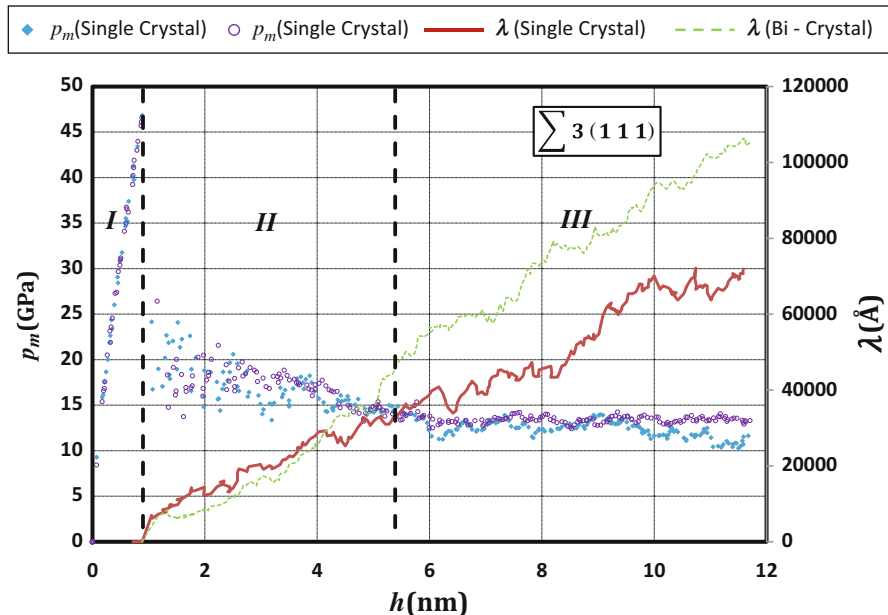


Fig. 20 Variation of mean contact pressure p_m and dislocation length λ as a function of indentation depth h for S2 bicrystal sample with $\Sigma 3(111)$ GB and its related single crystal sample (Reprinted with permission from Voyiadjis and Yaghoobi 2016)

length and consequently the hardness. The dominance of source exhaustion decreases during nanoindentation as more dislocations are provided to sustain the imposed deformation. Accordingly, the hardness reduction slope decreases during nanoindentation. Eventually, the dislocations reach the GB which blocks the dislocations.

- *Region III*: Enough dislocation length is provided to sustain the imposed deformation, and the source exhaustion hardening is not active anymore. The interactions of dislocations with each other and GB become important by increasing the dislocation content. Also, the number of dislocations blocked by GB becomes considerable and the produced pile-up enhances the sample strength. Consequently, the GB enhances the nanoindentation response of thin film for S2 sample.

The dislocation visualization of the S2 thin film with and without CTB is illustrated in Fig. 21 during nanoindentation. Figure 21a, b illustrates that the initial dislocation is homogeneously nucleated beneath the indenter which is a Shockley partial dislocation with the Burgers vector of $\frac{1}{6} [2\bar{1}1]$. The results show that the GB does not change the nucleation pattern. After the initial nucleation, Fig. 21c, d shows that the cross-slip is the controlling mechanism of deformation which increases the number of dislocation sources and provides the required dislocation content.

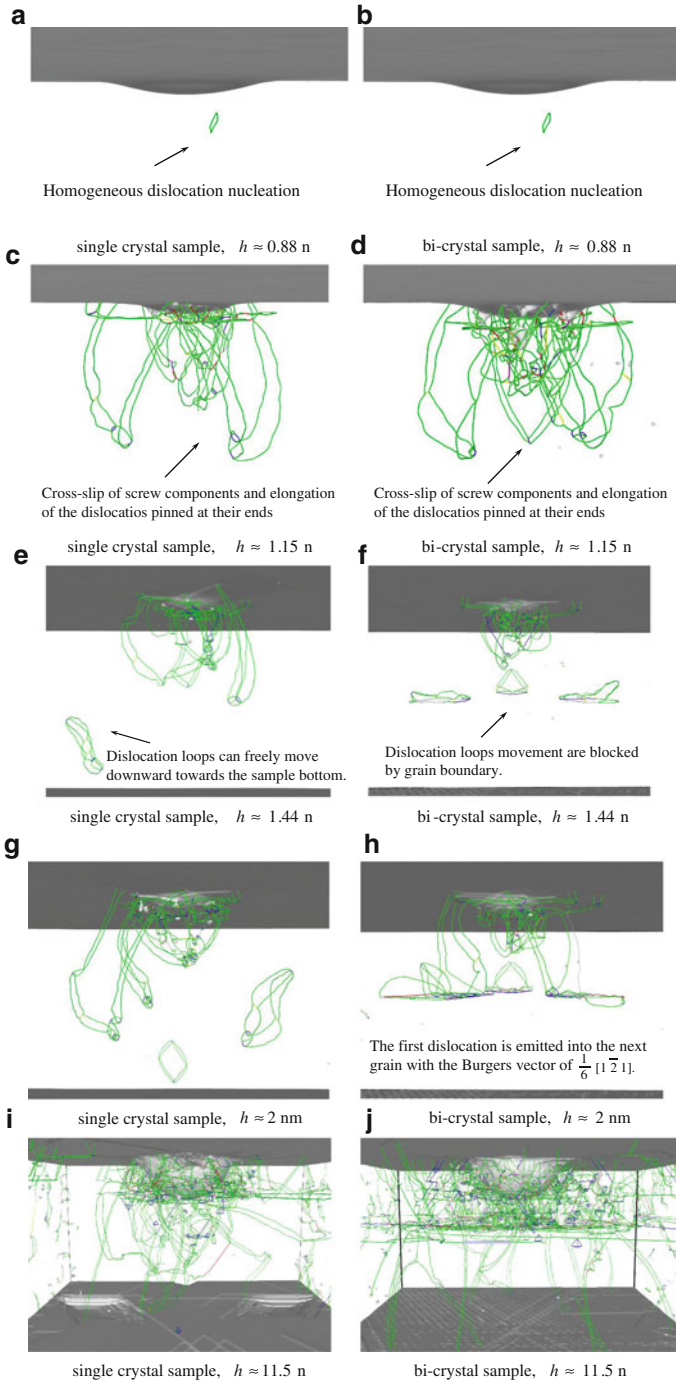


Fig. 21 (continued)

The effects of GB is still negligible on the dislocation pattern. The dislocation loops are induced by cross-slipping and pinching off of screw dislocations as the indentation depth increases. The induced loops are moving downward which are blocked by GB. Consequently, the GB starts to change the pattern of dislocation evolution. Figure 21e, f illustrates the dislocation loops movement in thin films with and without GB, respectively. The dislocation blocked by GB are eventually emitting into the next grain by increasing the indentation depth. Figure 21h illustrates the first dislocation emitting into the next grain which is a Shockley partial dislocation with the Burgers vector of $\frac{1}{6} [121]$. Figure 21i, j shows the dislocation visualization of the sample at the higher indentation depths for thin film with and without GB, respectively. Although some dislocations are emitted into the next grain, the visualization results show a considerable pile-up behind the GB, while the dislocations are moving downward freely for single crystal thin film.

Figure 22 shows the variation of p_m and λ during nanoindentation for the S2 samples with different GBs and their related single crystal thin films. The observed microstructural behavior for CTB can be incorporated for all other GBs except $\sum 11(332)$ and $\sum 11(225)/(441)$ GBs. In the cases of two latter GBs, the GB enhances the hardness while the total dislocation length of thin film with GB is very close to the one without GB. The observed discrepancy is due to the fact the total dislocation length is not an appropriate factor to study the forest hardening mechanism. In the case of source exhaustion hardening, the total dislocation length dictates the amount of stress required to sustain the plastic flow. On the other hand, the density of dislocation in the plastic zone should be taken as the representative factor for the forest hardening mechanism. Voyiadjis and Yaghoobi (2016) assumed that the plastic zone is located in the upper grain. Accordingly, the total dislocation length in the upper grain λ_{upper} should be investigated during nanoindentation. The dislocations located in the upper one third is considered for single crystal thin film, and the obtained results are compared with those of bicrystal thin films. Figure 23 compares the variations of mean contact pressure and the total dislocation length located in the plastic zone during nanoindentation for thin films with and without GB. The results show that the GB increases the total dislocation length located in plastic zone and consequently enhances the hardness according to the forest hardening mechanism. The results show that the main role of GB in the cases of large thin films, i.e., S2 samples, is to modify the pattern of dislocation in a way that increases the dislocation density located in the



Fig. 21 Dislocation nucleation and evolution: **a** single crystal sample, $h \approx 0.88$ nm; **b** bicrystal sample, $h \approx 0.88$ nm; **c** single crystal sample, $h \approx 1.15$ nm; **d** bicrystal sample, $h \approx 1.15$ nm; **e** single crystal sample, $h \approx 1.44$ nm; **f** bicrystal sample, $h \approx 1.44$ nm; **g** single crystal sample, $h \approx 2.03$ nm; **h** bicrystal sample, $h \approx 2.03$ nm; **i** single crystal sample, $h \approx 11.5$ nm; **j** bicrystal sample, $h \approx 11.5$ nm (Reprinted with permission from Voyiadjis and Yaghoobi 2016)

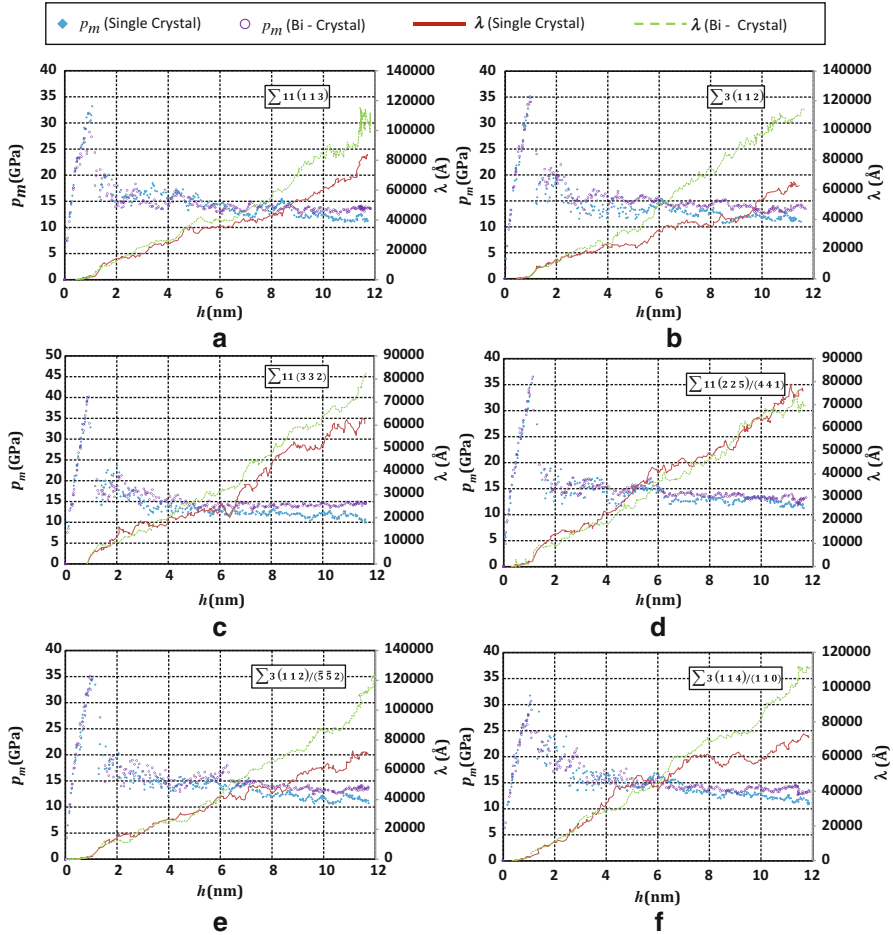


Fig. 22 Variation of mean contact pressure p_m and dislocation length λ as a function of indentation depth h for S2 bicrystal and their related single crystal samples with grain boundaries of: **a** $\sum 11(113)$, **b** $\sum 3(112)$, **c** $\sum 11(332)$, **d** $\sum 11(225)/(441)$, **e** $\sum 3(112)/(552)$, and **f** $\sum 3(114)/(110)$ (Reprinted with permission from Voyiadjis and Yaghoobi 2016)

plastic zone and accordingly strengthen the thin films. One should note that the strain rates incorporated in the atomistic simulation are much higher than those selected for experiments. Accordingly, the interpretation of the obtained results should be carefully handled. The applied strain rate can influence both hardening mechanisms and dislocation network properties (see, e.g., Yaghoobi and Voyiadjis, 2016b; Voyiadjis and Yaghoobi, 2017; Yaghoobi and Voyiadjis, 2017). In other words, one should ensure that the observed mechanisms are not artifacts of the high strain rates used in the atomistic simulation.

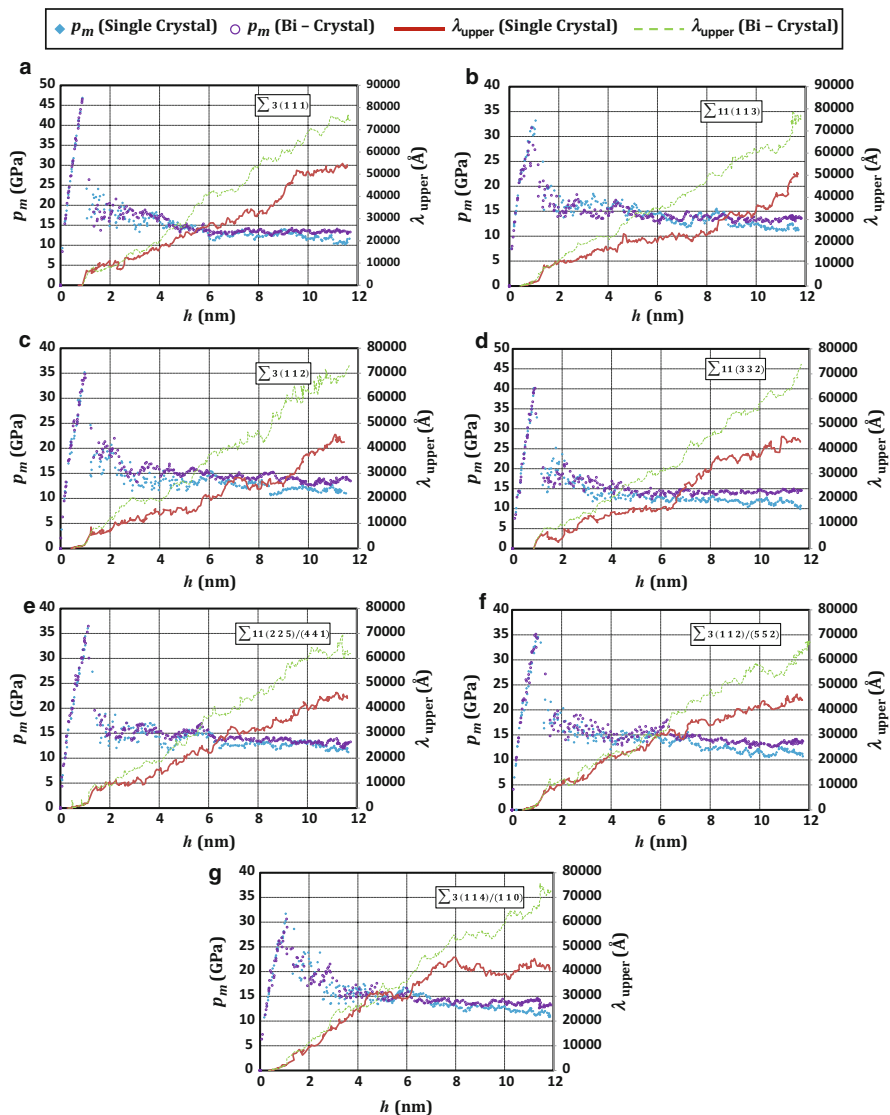


Fig. 23 Variation of mean contact pressure p_m and total dislocation length located in the upper grain λ_{upper} as a function of indentation depth h for S2 bicrystal and their related single crystal samples with grain boundaries of **a** $\Sigma 3(111)$, **b** $\Sigma 11(113)$, **c** $\Sigma 3(112)$, **d** $\Sigma 11(332)$, **e** $\Sigma 11(225)/(441)$, **f** $\Sigma 3(112)/(552)$, and **g** $\Sigma 3(114)/(110)$ (Reprinted with permission from Voyiadjis and Yaghoobi 2016)

References

- A.H. Almasri, G.Z. Voyiadjis, Nano-indentation in FCC metals: experimental study. *Acta Mech.* **209**, 1–9 (2010)
- R.K.A. Al-Rub, G.Z. Voyiadjis, Analytical and experimental determination of the material intrinsic length scale of strain gradient plasticity theory from micro- and nano-indentation experiments. *Int. J. Plast.* **20**, 1139–1182 (2004)
- M.I. Baskes, Modified embedded-atom potentials for cubic materials and impurities. *Phys. Rev. B* **46**, 2727 (1992)
- S.G. Corcoran, R.J. Colton, E.T. Lilleodden, W.W. Gerberich, Anomalous plastic deformation at surfaces: nanoindentation of gold single crystals. *Phys. Rev. B* **55**, 16057–16060 (1997)
- C.F.O. Dahlberg, Y. Saito, M.S. Öztop, J.W. Kysar, Geometrically necessary dislocation density measurements associated with different angles of indentations. *Int. J. Plast.* **54**, 81–95 (2014)
- M.S. Daw, M.I. Baskes, Embedded-atom method: derivation and application to impurities, surfaces, and other defects in metals. *Phys. Rev. B* **29**, 6443–6453 (1984)
- E. Demir, D. Raabe, N. Zaafarani, S. Zaeferrer, Investigation of the indentation size effect through the measurement of the geometrically necessary dislocations beneath small indents of different depths using EBSD tomography. *Acta Mater.* **57**, 559–569 (2009)
- E. Demir, D. Raabe, F. Roters, The mechanical size effect as a mean-field breakdown phenomenon: example of microscale single crystal beam bending. *Acta Mater.* **58**, 1876–1886 (2010)
- K. Durst, B. Backes, M. Göken, Indentation size effect in metallic materials: correcting for the size of the plastic zone. *Scr. Mater.* **52**, 1093–1097 (2005)
- J.A. El-Awady, Unravelling the physics of size-dependent dislocation-mediated plasticity. *Nat. Commun.* **6**, 5926 (2015)
- J.A. El-Awady, M. Wen, N.M. Ghoniem, The role of the weakest-link mechanism in controlling the plasticity of micropillars. *J. Mech. Phys. Solids* **57**, 32–50 (2009)
- D. Faken, H. Jonsson, Systematic analysis of local atomic structure combined with 3D computer graphics. *Comput. Mater. Sci.* **2**, 279–286 (1994)
- J.R. Greer, *Nano and Cell Mechanics: Fundamentals and Frontiers*. Wiley, Chichester, pp 163–190 (2013)
- A. Hasnaoui, P.M. Derlet, H. Van Swygenhoven, Interaction between dislocations and grain boundaries under an indenter – a molecular dynamics simulation. *Acta Mater.* **52**, 2251–2258 (2004)
- H. Jang, D. Farkas, Interaction of lattice dislocations with a grain boundary during nanoindentation simulation. *Mater. Lett.* **61**, 868–871 (2007)
- C.L. Kelchner, S.J. Plimpton, J.C. Hamilton, Dislocation nucleation and defect structure during surface indentation. *Phys. Rev. B* **58**, 11085–11088 (1998)
- C.C. Koch, I.A. Ovid'ko, S. Seal, S. Veprek, *Structural Nanocrystalline Materials: Fundamentals and Applications* (Cambridge University Press, Cambridge, 2007)
- M. de Koning, R.J. Kurtz, V.V. Bulatov, C.H. Henager, R.G. Hoagland, W. Cai, M. Nomura, Modeling of dislocation–grain boundary interactions in FCC metals. *J. Nucl. Mater.* **323**, 281–289 (2003)
- O. Kraft, P. Gruber, R. Mönig, D. Weygand, Plasticity in confined dimensions. *Annu. Rev. Mater. Res.* **40**, 293–317 (2010)
- Y. Kulkarni, R.J. Asaro, D. Farkas, Are nanotwinned structures in fcc metals optimal for strength, ductility and grain stability? *Scr. Mater.* **60**, 532–535 (2009)
- J.W. Kysar, C.L. Briant, Crack tip deformation fields in ductile single crystals. *Acta Mater.* **50**, 2367–2380 (2002)
- J.W. Kysar, Y.X. Gan, T.L. Morse, X. Chen, M.E. Jones, High strain gradient plasticity associated with wedge indentation into face-centered cubic single crystals: geometrically necessary dislocation densities. *J. Mech. Phys. Solids* **55**, 1554–1573 (2007)
- Y. Lee, J.Y. Park, S.Y. Kim, S. Jun, Atomistic simulations of incipient plasticity under Al (111) nanoindentation. *Mech. Mater.* **37**, 1035–1048 (2005)

- J. Li, K.J. Van Vliet, T. Zhu, S. Yip, S. Suresh, Atomistic mechanisms governing elastic limit and incipient plasticity in crystals. *Nature* **418**, 307–310 (2002)
- S.N. Medyanik, S. Shao, Strengthening effects of coherent interfaces in nanoscale metallic bilayers. *Comput. Mater. Sci.* **45**, 1129–1133 (2009)
- M.A. Meyers, A. Mishra, D.J. Benson, Mechanical properties of nanocrystalline materials. *Prog. Mater. Sci.* **51**, 427–556 (2006)
- Y. Mishin, D. Farkas, M.J. Mehl, D.A. Papaconstantopoulos, Interatomic potentials for monoatomic metals from experimental data and ab initio calculations. *Phys. Rev. B* **59**, 3393–3407 (1999)
- A.K. Nair, E. Parker, P. Gaudreau, D. Farkas, R.D. Kriz, Size effects in indentation response of thin films at the nanoscale: a molecular dynamics study. *Int. J. Plast.* **24**, 2016–2031 (2008)
- W.D. Nix, H.J. Gao, Indentation size effects in crystalline materials: a law for strain gradient plasticity. *J. Mech. Phys. Solids* **46**, 411–425 (1998)
- D.M. Norfleet, D.M. Dimiduk, S.J. Polasik, M.D. Uchic, M.J. Mills, Dislocation structures and their relationship to strength in deformed nickel microcrystals. *Acta Mater.* **56**, 2988–3001 (2008)
- T.A. Parthasarathy, S.I. Rao, D.M. Dimiduk, M.D. Uchic, D.R. Trinkle, Contribution to size effect of yield strength from the stochastics of dislocation source lengths in finite samples. *Scr. Mater.* **56**, 313–316 (2007)
- P. Peng, G. Liao, T. Shi, Z. Tang, Y. Gao, Molecular dynamic simulations of nanoindentation in aluminum thin film on silicon substrate. *Appl. Surf. Sci.* **256**, 6284–6290 (2010)
- S. Plimpton, Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **117**, 1–19 (1995)
- N.M. Pugno, A general shape/size-effect law for nanoindentation. *Acta Mater.* **55**, 1947–1953 (2007)
- S.I. Rao, D.M. Dimiduk, M. Tang, T.A. Parthasarathy, M.D. Uchic, C. Woodward, Estimating the strength of single-ended dislocation sources in micron-sized single crystals. *Philos. Mag.* **87**, 4777–4794 (2007)
- S.I. Rao, D.M. Dimiduk, T.A. Parthasarathy, M.D. Uchic, M. Tang, C. Woodward, Athermal mechanisms of size-dependent crystal flow gleaned from three-dimensional discrete dislocation simulations. *Acta Mater.* **56**, 3245–3259 (2008)
- M.D. Sangid, T. Ezaz, H. Sehitoglu, I.M. Robertson, Energy of slip transmission and nucleation at grain boundaries. *Acta Mater.* **59**, 283–296 (2011)
- S. Shao, S.N. Medyanik, Dislocation–interface interaction in nanoscale fcc metallic bilayers. *Mech. Res. Commun.* **37**, 315–319 (2010)
- W.A. Soer, J.T.M. De Hosson, Detection of grain-boundary resistance to slip transfer using nanoindentation. *Mater. Lett.* **59**, 3192–3195 (2005)
- A. Stukowski, Structure identification methods for atomistic simulations of crystalline materials. *Model. Simul. Mater. Sci. Eng.* **20**, 045021 (2012)
- A. Stukowski, Computational analysis methods in atomistic modeling of crystals. *JOM* **66**, 399–407 (2014)
- A. Stukowski, K. Albe, Extracting dislocations and non-dislocation crystal defects from atomistic simulation data. *Model. Simul. Mater. Sci. Eng.* **18**, 085001 (2010)
- A. Stukowski, K. Albe, D. Farkas, Nanotwinned fcc metals: strengthening versus softening mechanisms. *Phys. Rev. B* **82**, 224103 (2010)
- A. Stukowski, V.V. Bulatov, A. Arsenlis, Automated identification and indexing of dislocations in crystal interfaces. *Model. Simul. Mater. Sci. Eng.* **20**, 085007 (2012)
- S. Suresh, T.G. Nieh, B.W. Choi, Nanoindentation of copper thin films on silicon substrates. *Scr. Mater.* **41**, 951–957 (1999)
- J.G. Swadener, E.P. George, G.M. Pharr, The correlation of the indentation size effect measured with indenters of various shapes. *J. Mech. Phys. Solids* **50**, 681–694 (2002)
- J. Tersoff, New empirical approach for the structure and energy of covalent systems. *Phys. Rev. B* **37**, 6991–7000 (1988)

- T. Tsuru, Y. Kaji, D. Matsunaka, Y. Shibutani, Incipient plasticity of twin and stable/unstable grain boundaries during nanoindentation in copper. *Phys. Rev. B* **82**, 024101 (2010)
- M.D. Uchic, P.A. Shade, D.M. Dimiduk, Plasticity of micrometer-scale single crystals in compression. *Annu. Rev. Mater. Res.* **39**, 361–386 (2009)
- G.Z. Voyiadjis, R.K.A. Al-Rub, Gradient plasticity theory with a variable length scale parameter. *Int. J. Solids Struct.* **42**, 3998–4029 (2005)
- G.Z. Voyiadjis, M. Yaghoobi, Large scale atomistic simulation of size effects during nanoindentation: dislocation length and hardness. *Mater. Sci. Eng. A* **634**, 20–31 (2015)
- G.Z. Voyiadjis, M. Yaghoobi, Role of grain boundary on the sources of size effects. *Comput. Mater. Sci.* **117**, 315–329 (2016)
- G.Z. Voyiadjis, M. Yaghoobi, Size and strain rate effects in metallic samples of confined volumes: dislocation length distribution. *Scr. Mater.* **130**, 182–186 (2017)
- M. Yaghoobi, G.Z. Voyiadjis, Effect of boundary conditions on the MD simulation of nanoindentation. *Comput. Mater. Sci.* **95**, 626–636 (2014)
- M. Yaghoobi, G.Z. Voyiadjis, Atomistic simulation of size effects in single-crystalline metals of confined volumes during nanoindentation. *Comput. Mater. Sci.* **111**, 64–73 (2016a)
- M. Yaghoobi, G.Z. Voyiadjis, Size effects in fcc crystals during the high rate compression test. *Acta Mater.* **121**, 190–201 (2016b)
- M. Yaghoobi, G.Z. Voyiadjis, Microstructural investigation of the hardening mechanism in fcc crystals during high rate deformations. *Comp. Mater. Sci.* **138**, 10–15 (2017)
- N. Zaafarani, D. Raabe, F. Roters, S. Zaeferrer, On the origin of deformation-induced rotation patterns below nanoindents. *Acta Mater.* **56**, 31–42 (2008)
- T.T. Zhu, A.J. Bushby, D.J. Dunstan, Materials mechanical size effects: a review. *Mater. Technol.* **23**, 193–209 (2008)
- J.A. Zimmerman, C.L. Kelchner, P.A. Klein, J.C. Hamilton, S.M. Foiles, Surface step effects on nanoindentation. *Phys. Rev. Lett.* **87**, 165507 (2001)



Molecular Dynamics-Decorated Finite Element Method (MDeFEM): Application to the Gating Mechanism of Mechanosensitive Channels

3

Liangliang Zhu, Qiang Cui, Yilun Liu, Yuan Yan, Hang Xiao,
and Xi Chen

Contents

Introduction	79
A Brief Review of MscS Studies	79
A Brief Review of MscL Studies	81
Molecular Dynamics-Decorated Finite Element Method (MDeFEM)	82
Gating and Inactivation of <i>E. coli</i> -MscS	83
Models and Methods	83
Results and Discussion	89

L. Zhu

Columbia Nanomechanics Research Center, Department of Earth and Environmental Engineering, Columbia University, New York, NY, USA

International Center for Applied Mechanics, State Key Laboratory for Strength and Vibration of Mechanical Structures, School of Aerospace, Xi'an Jiaotong University, Xi'an, China
e-mail: seven.zhu@qq.com; zhu.liangliang@stu.xjtu.edu.cn

Q. Cui

Department of Chemistry and Theoretical Chemistry Institute, University of Wisconsin-Madison, Madison, WI, USA
e-mail: cui@chem.wisc.edu

Y. Liu

International Center for Applied Mechanics, State Key Laboratory for Strength and Vibration of Mechanical Structures, School of Aerospace, Xi'an Jiaotong University, Xi'an, China
e-mail: yilunliu@mail.xjtu.edu.cn

Y. Yan · H. Xiao

School of Chemical Engineering, Northwest University, Xi'an, China
e-mail: deeplake@qq.com; xiaohang007@gmail.com

X. Chen (✉)

Department of Earth and Environmental Engineering, Columbia Nanomechanics Research Center, Columbia University, New York, NY, USA
e-mail: xichen@columbia.edu

Coupled Continuum Mechanical-Continuum Solvation Approach with Application to Gating Mechanism of MscL	99
Models and Methods	99
Results and Discussion	110
Limitations of the Current Implementation and Future Directions	122
Concluding Remarks	124
References	125

Abstract

Many fundamentally important biological processes rely on the mechanical responses of membrane proteins and their assemblies in the membrane environment, which are multiscale in nature and represent a significant challenge in modeling and simulation. For example, in mechanotransduction, mechanical stimuli can be introduced through macroscopic-scale contacts, which are transduced to mesoscopic-scale (micron) distances and can eventually lead to microscopic-scale (nanometer) conformational changes in membrane-bound protein or protein complexes. This is a fascinating process that spans a large range of length scales and time scales. The involvement of membrane environment and critical issues such as cooperativity calls for the need for an efficient multi-scale computational approach. The goal of the present research is to develop a hierarchical approach to study the mechanical behaviors of membrane proteins with a special emphasis on the gating mechanisms of mechanosensitive (MS) channels. This requires the formulation of modeling and numerical methods that can effectively bridge the disparate length and time scales. A top-down approach is proposed to achieve this by effectively treating biomolecules and their assemblies as integrated structures, in which the most important components of the biomolecule (e.g., MS channel) are modeled as continuum objects, yet their mechanical/physical properties, as well as their interactions, are derived from atomistic simulations. Molecular dynamics (MD) simulations at the nanoscale are used to obtain information on the physical properties and interactions among protein, lipid membrane, and solvent molecules, as well as relevant energetic and temporal characteristics. Effective continuum models are developed to incorporate these atomistic features, and the conformational response of macromolecule(s) to external mechanical perturbations is simulated using finite element (FEM) analyses with in situ mechanochemical coupling. Results from the continuum mechanics analysis provide further insights into the gating transition of MS channels at structural and physical levels, and specific predictions are proposed for further experimental investigations. It is anticipated that the hierarchical framework is uniquely suited for the analysis of many biomolecules and their assemblies under external mechanical stimuli.

Keywords

Mechanotransduction · Multi-scale simulation · Mechanosensitive channels · Gating mechanism · Continuum mechanics · Continuum solvation

Introduction

Occurring over large time and length scales, various biological signal transduction processes rely on the mechanical response of biomolecules and their assemblies to external stimuli. Muscle contraction or stretch, as a prominent example, involves the cooperative mechanical response of a large number of myosin molecules (Geeves and Holmes 1999, 2005), and structural changes from molecular scale to organ scales of muscle contribute greatly to its various and remarkable adaptations under different mechanical stimuli (Wisdom et al. 2014). Another example is mechanosensation (Hamill and Martinac 2001), during which the mechanosensitive (MS) channels play important roles in living cells of diverse phylogenetic origin (Martinac 2004) and have been identified in more than 30 cell types (Sackin 1995). By converting mechanical forces exerted on the cell membrane into biochemical or electrical signals, MS channels are involved in a wide range of cellular processes including cell growth and differentiation (Wang and Thampatty 2006) and blood pressure and cell volume regulation (Hamill and Martinac 2001; Martinac and Kloda 2012; Sun et al. 2009) and are essential to sensations such as touching, balance, and hearing (Hamill and Martinac 2001; Ingber 2006; Martinac 2004). A direct link between the lipid membrane and the structure/function of some MS channels has been revealed (Phillips et al. 2009). And in eukaryotic cells, the cytoskeleton was shown to play a similar role in the activation of MS channels (Hayakawa et al. 2008). With an increasing number of MS channels being identified, their atomic structures, gating characteristics, and functional mechanisms have been studied extensively in the past decades. Among the families of MS channels, a much-studied system is the mechanosensitive channel of small/large conductance (MscS/MscL) in *Escherichia coli* (*E. coli*), which serves as a paradigm for understanding the gating behaviors of the MS family of ion channels. Functioning as the “safety valve” of bacteria that regulates turgor pressure, MscS/MscL is sensitive to tension in the membrane, and the opening of MscS/MscL allows exchange of ions (nonselective between anions and cations) and small molecules (including water) between the cytoplasm and the environment (Berrier et al. 1996; Blount et al. 1997; Martinac et al. 2014; Saimi et al. 1992).

A Brief Review of MscS Studies

The first crystal structure for *E. coli*-MscS was solved at 3.95 Å resolution (Bass et al. 2002) and subsequently refined to a higher resolution of 3.7 Å (Steinbacher et al. 2007). The structure has a pore of less than 5 Å in diameter, and because of its hydrophobic constriction, the pore is thus considered as nonconducting or closed (Steinbacher et al. 2007; Vora et al. 2006). The open form of the A106V mutant of MscS was subsequently crystallized at 3.45 Å resolution (Wang et al. 2008). The crystal structures of *E. coli*-MscS have been constantly challenged, mainly due to the large voids between the TM3 helix (the third transmembrane helix) and

the closely packed TM1 and TM2 helices (also referred as TM pockets). Using extrapolated motion dynamics (EMD) (Akitake et al. 2007) and continuous wave electron paramagnetic resonance (cwEPR) (Vasquez et al. 2008a, b), alternative MscS structures of both closed and open forms were independently generated, in which the apparent voids in the crystal structures were absent. The three approaches lead to three sets of mutually incompatible models of the closed and open structures and thus models for the gating transition. A pulsed electron-electron double resonance (PELDOR) approach reevaluated these competing structural models both in detergent (Pliotas et al. 2012) and in bilayer mimics (Ward et al. 2014); the results supported the arrangement of helices seen in the crystal structures. Another study reported the crystal structure of β -dodecylmaltoside-solubilized wild-type *E. coli*-MscS at 4.4 Å resolution and further supported that the A106V structure resembles the open state (Lai et al. 2013). Finally, a higher-resolution structure of the *E. coli*-MscS identified alkyl chains inside the pockets/voids formed by the transmembrane helices (Pliotas et al. 2015), strongly support that the voids in *E. coli*-MscS crystal structures are realistic (Pliotas and Naismith 2016). Based on the above evidence, the present study starts with the assumption that the crystal structure at 3.7 Å resolution (Steinbacher et al. 2007) and the A106V mutant at 3.45 Å resolution (Wang et al. 2008) represent the closed (resting) and open states of MscS, respectively, but noting that the specific functional states of these structures of MscS may be still in debate. Nevertheless, despite the available closed and open structures, little is known about the partially expanded intermediate structures during MscS gating transition.

Numerous studies have explored residues and interactions that are important to the gating characteristics of MscS. Some of the established cases include the Asp62-Arg131/Arg128 salt bridges (Nomura et al. 2008) and the Phe68-Leu111/Leu115 (Belyy et al. 2010) apolar interaction, which affect channel gating and inactivation; Leu105 and Leu109 form a hydrophobic lock at the channel pore (Anishkin and Sukharev 2004; Vora et al. 2006); the interaction between the lower part of TM3 and the cytoplasmic β domain and Gly113 is crucial to inactivation (Edwards et al. 2008; Koprowski et al. 2011; Petrov et al. 2013); and a number of residues were shown to influence force transmission at the protein-lipid interface (Malcolm et al. 2011; Nomura et al. 2006). The physical origins for the importance of these interactions are not always well understood. Another much studied mechanistic issue concerns inactivation of MscS: under prolonged exposure to subthreshold membrane tension, the channel desensitizes into an inactivated and nonconducting state from which it must relax back to the closed state in lower membrane tension before reactivation can be induced (Akitake et al. 2005; Edwards et al. 2008; Koprowski and Kubalski 1998; Levina et al. 1999). Up to now, while multiple residues and interactions are known to be important to inactivation (Vasquez 2013), the structural mechanism underlying inactivation remains elusive.

Furthermore, it has been reported that the open probability of MscS can be significantly increased by membrane depolarization (Cui et al. 1995; Martinac et al. 1987). A later study, however, showed that the activation of *E. coli*-MscS by membrane tension is essentially independent of the transmembrane voltage (Akitake et al. 2005), though depolarizing membrane voltage strongly promotes inactivation

(Akitake et al. 2005). In addition, a recent study reported that the arginine residues at positions of 46, 54, and 74 in TM1 and TM2 helices are not responsible for the voltage dependence of inactivation (Nomura et al. 2016). The structural response and inactivation mechanism under membrane potential thus remain to be better clarified.

Atomistic simulations have made valuable contributions to the understanding of ion channels, including MscS, in recent years (Akitake et al. 2007; Anishkin and Sukharev 2004; Deplazes et al. 2012; Masetti et al. 2016; Pliotas et al. 2015; Sotomayor and Schulten 2004; Sotomayor et al. 2006). Nevertheless, such simulations remain computationally intensive, making it difficult to study gating transitions that occur on the millisecond time scale and explore contributions of specific structural motifs and interactions.

A Brief Review of MscL Studies

A large body of experimental, theoretical, and simulation work has focused on elucidating the molecular mechanism of MscL gating (Booth et al. 2007; Haswell et al. 2011). Experiments directly probing the gating transition of MscL were primarily patch-clamp measurements, which simultaneously monitor the membrane tension and ionic currents through the channel (thus opening probability) (Sukharev et al. 1997, 1999). An important clue from these studies on MscL reconstituted into purified lipid bilayers is that the mechanical property of the membrane plays a principal role in determining the gating behaviors of MscL. A model with five subconducting states was established (Sukharev et al. 1999), in which the tension-dependent conformational transition was primarily attributed to the pore area variation that occurs between the closed state and a low subconductance state. In addition, other experimental studies have been used to probe MscL's conformational transition, and these include electron paramagnetic resonance spectroscopy (EPR) with site-directed spin labeling (SDSL) (Perozo et al. 2002a, b) for monitoring the structural rearrangements, cysteine scanning for identifying residue contacts in the transmembrane helices (Levin and Blount 2004), and numerous mutation studies for probing the importance of residues in different structural motifs (Anishkin et al. 2005; Blount et al. 1997; Levin and Blount 2004; Tsai et al. 2005). Besides, a single-molecule fluorescence resonance energy transfer (FRET) method (Wang et al. 2014) or the combination of data from FRET spectroscopy and simulations (Corry et al. 2010) has enabled a more detailed description of the open form of MscL in the natural lipid environment. Based on geometrical constraints provided by various measurements and the crystal structure of MscL in *Mycobacterium tuberculosis* (*Tb*) (Chang et al. 1998), structural models for the closed-open transition of *E. coli*-MscL were constructed (Sukharev et al. 2001a, b); revised models were proposed subsequently where the conformational changes of the S3 helices are much smaller in scale (Sukharev and Anishkin 2004; Sukharev and Corey 2004). Although highly valuable, these structural models need to be evaluated for validity in a systematic and physical manner (Chen et al. 2008; Tang et al. 2006, 2008).

Meanwhile, analytical models for the gating transition in MscL have been developed by several groups. Markin and Sachs presented a general thermodynamic

model for mechanotransduction that relates the probability of channel opening to membrane properties such as thickness, curvature, and stiffness (Markin and Sachs 2004). An analytical continuum model was developed by Wiggins and Phillips to characterize the free energy of the protein-bilayer system (Wiggins and Phillips 2004); the model highlighted that the competition of hydrophobic mismatch could be a physical gating mechanism. As an alternative to dilatational gating, a gating-by-tilt model was proposed (Turner and Sens 2004) in which the gate opening is due to the swinging of the lipids near the channel with respect to a pivot. Although these analytical models are valuable for highlighting the potential contribution of specific physical factors (e.g., hydrophobic mismatch), they lack sufficient structural details to make specific connection with experimental studies.

Numerical simulation is a powerful approach for exploring the fundamental principles of mechanotransduction. To properly assign structural features to important intermediate states along the closed-open transition, it is important to simulate the structural response of the channel to mechanical perturbation consistent with the experimental protocol. Due to the large length scale and time scale involved, however, this is usually beyond the capability of atomistic simulations despite the rapid progresses being made (Deplazes et al. 2012; Dror et al. 2012; Gullingsrud and Schulten 2003; Karplus and Kuriyan 2005; Klepeis et al. 2009; Monticelli et al. 2008; Sawada et al. 2012; Snow et al. 2005; Yefimov et al. 2008). Hence, developing coarse-grained models to access longer time scales has become an important topic in the simulation community (Ingolfsson et al. 2014; Marrink et al. 2007; Marrink and Tieleman 2013; Monticelli et al. 2008; Praprotnik et al. 2008; Saunders and Voth 2012; Shi et al. 2006; Shinoda et al. 2012; Yefimov et al. 2008). Most of these efforts have focused on developing particle-based models in which one bead represents a group of atoms. Specifically for MscL, building upon their success in developing an effective coarse-grained model for lipids, Marrink et al. developed a coarse-grained model (Yefimov et al. 2008) for MscL based on the transfer free energy of amino acids between water and lipids. The gating transition was successfully observed in the simulation although the pore radius in the final state is somewhat smaller than that estimated in the literature (Sukharev et al. 2001b).

Molecular Dynamics-Decorated Finite Element Method (MDeFEM)

In light of the limitations of previous experimental/theoretical/numerical efforts concerning the structural rearrangements of MS channels during gating, it is worthwhile pursuing the alternative approach of continuum mechanical modeling, which has been used in a broad set of mechanics problems (Scarpa et al. 2010; Tang et al. 2006; Tserpes and Papanikos 2009; Zeng et al. 2012). Along this line, the establishment of molecular dynamics-decorated finite element method (MDeFEM) (Chen et al. 2008; Tang et al. 2006, 2008) represents an attractive alternative to atomistic and particle-based coarse-grained simulations, allowing the analysis of bimolecular systems at long time scales while maintaining sufficient molecular details to capture the most essential characteristics of the system under study.

In MDeFEM, biomolecules are modeled as integrated continuum motifs and the finite element simulation framework allows efficient treatment of deformations at large length scales and complex deformation modes inaccessible to conventional all-atom simulations.

In this chapter, MDeFEM (Chen et al. 2008; Tang et al. 2006, 2008) is firstly adapted to study the gating mechanism of MscS. The high computational efficiency of MDeFEM allows us to analyze the contributions of various structural motifs and interactions to the gating process as well as the effect of voltage. The observation of different gating characteristics upon perturbation of material properties of the helical kink region in TM3 at Gly113 also leads to the proposal of a mechanism for inactivation. Overall, the current simulations not only provide new insights into the gating transition of MscS with structural details but also lead to specific predictions that can be tested by future experimental studies.

Secondly, a number of major limitations of previous MDeFEM models of MscL (Chen et al. 2008; Tang et al. 2006, 2008) (or the MDeFEM models of MscS) need to be alleviated for more quantitative analysis. These limitations include (1) not sufficient structural/energetic details of MscL. The helices were represented by rounded sticks, and the inter-component interactions were computed based on surface-to-surface interactions; the nonbonded interactions among loops and those between the loops and the helices were not considered. (2) The lipid bilayer was treated as an elastic solid slab, whereas the realistic membrane should be fluidic and does not sustain large shear stress. (3) No solvation contribution was considered. It has been proposed that solvation plays a major role in stabilizing the open conformation of MscL due to the exposure of hydrophilic residues (Anishkin et al. 2005; Anishkin and Kung 2005). These limitations may have led to the exceedingly high membrane strain required for the full-gating transition of MscL in the previous studies (Chen et al. 2008; Tang et al. 2008). The MDeFEM approach is here improved significantly to address these issues through a coupled mechanical-chemical approach. More realistic models for the MscL molecule and the membrane are developed. To include solvation effects, a force-based protocol is established to integrate a continuum mechanics model for the mechanical properties of the macromolecule with a continuum treatment of solvation. A similar approach has been applied to study the salt concentration dependence of DNA bendability (Ma et al. 2009). It is envisioned that the high computational efficiency and flexibility will make this hierarchical multi-scale framework uniquely applicable to the study of mechanical behaviors of various biological systems, interpreting existing and stimulating new experimental investigations.

Gating and Inactivation of *E. coli*-MscS

Models and Methods

In MDeFEM, the protein structure is described by continuum FEM models, and nonbonded interactions between different components are represented by nonlinear

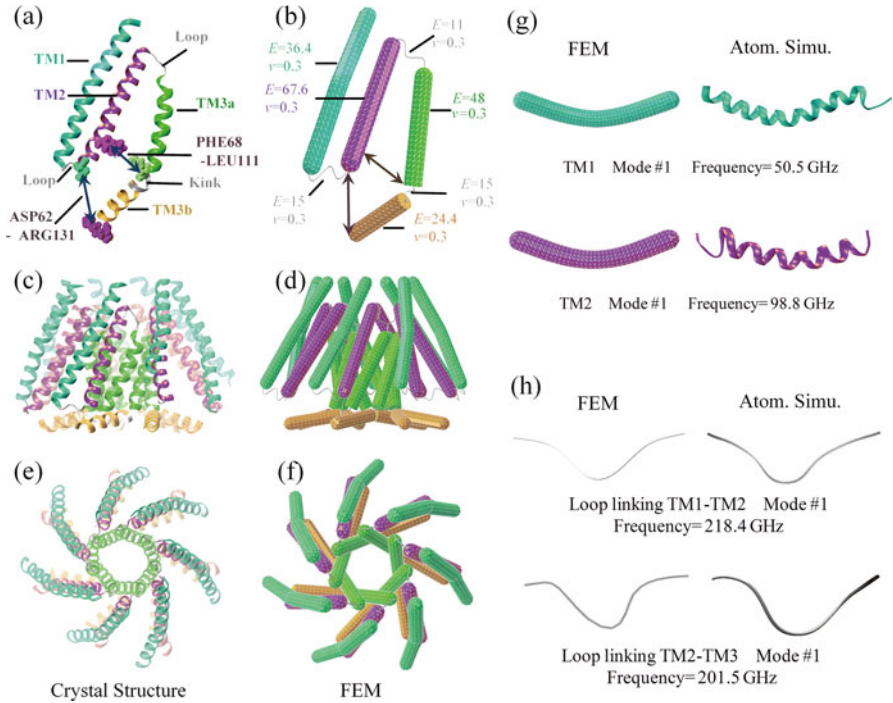


Fig. 1 Continuum mechanics modeling of MscS. **(a)** The backbone structure of one MscS monomer with two key interaction pairs (Asp62-Arg131 and Phe68-Leu111) highlighted by arrows. **(b)** The corresponding FEM model of one monomer. The Young's modulus, E (GPa), and Poisson's ratio, ν , of each component in one monomer are indicated aside. **(c–d)** Side views of the crystal structure and FEM model. **(e–f)** Top views of the crystal structure and FEM model. Examples of several lowest eigenmodes and frequencies of helices and loops are displayed in **(g)** and **(h)**, respectively. Results of molecular mechanics are compared to those of the finite element simulations

distance-dependent pressures or point-to-point connectors. In the following, the main continuum construction procedures of *E. coli*-MscS are introduced briefly with the commercial package ABAQUS. More details can be found in our previous works (Chen et al. 2008; Zhu et al. 2016). In the subsequent discussion, unless otherwise specified, MscS refers to *E. coli*-MscS.

A MscS channel (Bass et al. 2002) is formed by seven chains assembled as a heptamer with a large cytoplasmic region that exhibits a mixed α/β structure with several strands and α -helices intertwined together to form a balloon-like osmolyte filter cage (Gamini et al. 2011). In the transmembrane domain (Fig. 1c, e), each monomer (Fig. 1a) of MscS consists of three helices, referred to as TM1, TM2, and TM3, the last of which has a pronounced kink at the Gly113 region and is thus split into TM3a and TM3b. Each TM1 helix and its nearest neighbor, a TM2 helix, are assembled into a pair through a periplasmic loop, forming the outer boundary that interacts extensively with the lipid membrane; the TM3a helices form the inner

boundary for the pore, and TM3b helices splay out to be nearly parallel to the membrane plane.

Based on the closed backbone structure of MscS (Bass et al. 2002; Steinbacher et al. 2007), within the continuum mechanics framework, each helix is modeled as a three-dimensional elastic cylinder of 5 Å diameter and the loops as quasi-one-dimensional winding beams with cross-sectional diameter of 2.5 Å (Fig. 1b) (Tang et al. 2006; Zhu et al. 2016). For simplicity, only the transmembrane domain of MscS is considered (Anishkin et al. 2008b). Although there is evidence supporting that the cytoplasmic domain swells up (Machiyama et al. 2009) during gating, it is also suggested to be nonessential but only responsible for increased stability and activity (Schumann et al. 2004); the cytoplasmic domain is not expected to undergo large changes during the gating transition (Pliotas et al. 2012; Wang et al. 2008). The endpoints where the cytoplasmic domain is truncated are softly restrained in the continuum model (1 kcal/mol/Å^2) (Anishkin et al. 2008a, b; Spronk et al. 2006). Furthermore, the first 26 residues (in the N-termini) of the 286 amino acids of each MscS monomer were not resolved in the crystal structure (Bass et al. 2002) and therefore are also excluded in the continuum model; an experimental study showed that MscS can tolerate small deletions at the N-terminus (Miller et al. 2003).

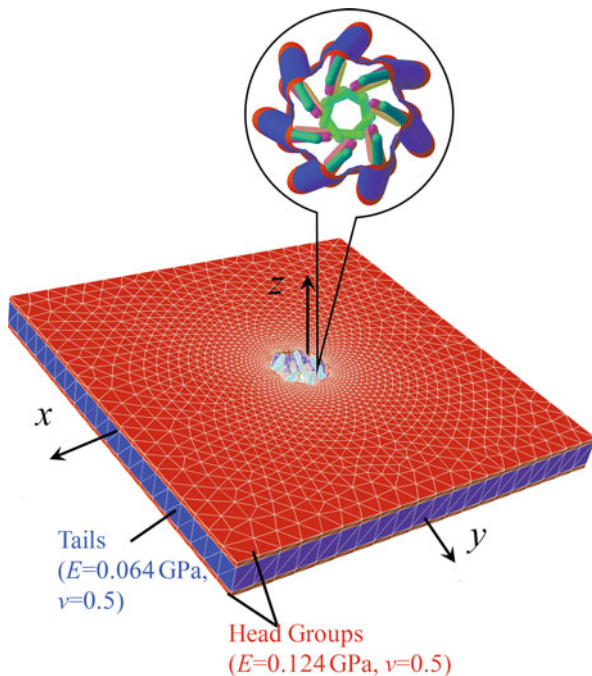
Material properties of each component of the continuum model are calibrated by matching results of normal mode analysis (NMA) at the atomistic and continuum levels (Fig. 1g–h). The key mechanical properties for the helices and loops are shown in Fig. 1b, and these are much larger than the estimated range of the Young's modulus of MscL α -helices (0.2 to 12.5 GPa) in Martinac et al.'s recent work (Bavi et al. 2017) with constant-force steered molecular dynamics (SMD). This is because, in the present study, the helices, for instance, are modeled as much thinner elastic cylinders (with a diameter of 5 Å) by considering only the main chain. The bending stiffness of the helices in the present continuum model and that of MscS helices by SMD are expected to be consistent. As a qualitative comparison, for example, the bending stiffness EI (the product of Young's modulus and moment of inertia) of MscS helices in vacuum is in the range of 70~200 (10^{-10} N Å^2) in this work, while based on Martinac et al.'s study (Bavi et al. 2017), the bending stiffness of MscL helices in water can be calculated to be in the range of 1~80 (10^{-10} N Å^2). For simplicity, the softening effect of the helices due to hydration (Anishkin and Sukharev 2017; Bavi et al. 2017) is not considered here, which may be one of the reasons for the difference in the above comparison of the bending stiffness. In both MscL (Bavi et al. 2017) and MscS, the second helix, TM2, is the stiffest one among the α -helices in each channel.

The MscS continuum model is embedded into an elastic membrane modeled as a sandwich panel that consists of three layers to mimic the lipid head and tail regions. A flat square membrane with a size of $400 \times 400 \text{ Å}$ is employed. To embed the channel (Fig. 1) into the continuum membrane, a cavity with the shape of a multi-petal flower (Fig. 2) is created in the middle of the membrane with the size and shape of the cavity conforming to those of MscS transmembrane helices in the closed state with an equilibrium distance of $\sim 5.5 \text{ Å}$ (Chen et al. 2008) from the surface of the cavity to the surfaces of the helices. The lipids buried in the voids

within the transmembrane helices (Pliotas et al. 2015) are neglected; their specific role in gating is not quite clear. The membrane model is parameterized based on MD calculations of the density map of water and lipids and the lateral pressure profiles of the POPE lipid membrane (Chen et al. 2008; Gullingsrud and Schulten 2004). Mechanical properties of the lipid membrane are shown in Fig. 2, and the in-plane shear modulus of the continuum membrane model is reduced (Zhu et al. 2016) to take into account the fluidity of the lipid membrane, i.e., its incapability of sustaining a large shear stress. It would be of interest to investigate the effect of the buried lipids (Pliotas et al. 2015) or more complex lipid properties (such as viscoelasticity (Deseri et al. 2016)) on gating, but we leave it to future studies since more sophisticated continuum models are required.

The assembled protein-membrane continuum model is shown in Fig. 2. Helices of MscS are meshed by four-node tetrahedron finite elements and loops by one-dimensional beam elements (Fig. 1). Without over-resolving the system, appropriate mesh density is ensured through convergence studies, and a typical element size of 1.5 Å is chosen. Four-node tetrahedron finite elements are also used to mesh the lipid membrane model with the mesh gradually more refined toward the boundary of the inner hole where it interacts with the protein directly (Fig. 2). To be consistent with the simple description of the protein-lipid continuum model, solvent molecules are not included, though the hydration effect can be further studied with a more sophisticated continuum mechanics-solvation coupled approach (Zhu et al. 2016) which is to be introduced in section “[Coupled Continuum Mechanical-Continuum Solvation Approach with Application to Gating Mechanism of MscL.](#)”

Fig. 2 Schematic of the continuum protein-lipid membrane model. A zoomed-in view is shown to illustrate the multi-petal lipid hole encompassing the MscS protein



In this work, we explicitly consider two key interaction pairs known to be essential to MscS gating, i.e., the Asp62-Arg131 salt bridge and the Phe68-Leu111 apolar interaction (Fig. 1a) mentioned in the section “Introduction.” The nonbonded interaction energy (electrostatic and van der Waals) between each pair of residues is calculated using the CHARMM 36 force field (Best et al. 2012; Brooks et al. 1983) as a function of distance along the center-of-mass separation vector (Fig. 3b). An electrostatic screening factor of 6 is used when calculating the electrostatic interaction for the Asp62-Arg131 salt bridge, considering that the gap between Asp62-Arg131 residues in the same monomer or across monomers is spanned mainly by protein atoms or lipid molecules rather than water (Pliotas et al. 2015); varying this factor from 3 to 10 generally has little impact on the computed gating behavior (data not show). The nonlinear distance-dependent interaction force between two residues in each pair is derived accordingly from the energy profile and applied to the closest finite element nodes by invisible nonlinear connector elements (arrows in Fig. 1b) in ABAQUS (Zhu et al. 2016). The connector element representing a pair of key interaction is defined both within the same monomer and across neighbor monomers.

Apart from the above two specific residue pairs, interactions between different helices and between the helix and the lipid membrane are described by a pair-wise effective pressure-distance relationship where the atom-to-atom interaction in the atomic structure are averaged to the surfaces of different components (Tang et al. 2006) (see Fig. 3a). The nonbonded (electrostatic and van der Waals) interactions are calculated without any cutoff. Incorporating an implicit membrane environment through the GBSW model (Im et al. 2003) in CHARMM in the inter-helical calculations has a negligible impact on the energy profiles. Taking the first derivative of the interaction energy with respect to distance, the pressure-distance relationship between two surfaces (adopting the sign convention that repulsive pressure is positive) is obtained and takes the following form:

$$p(\alpha_i) = \frac{Cn}{d_0} \left[(d_0/\alpha_i)^{m+1} - (d_0/\alpha_i)^{n+1} \right] \quad (1)$$

where p and d_0 are the interaction pressure and equilibrium distance between two surfaces, respectively, C is the energy well-depth, and α_i is the instantaneous distance between two deformed surfaces for the i -th element. This nonbonded pressure model implicitly includes both electrostatic and van der Waals (VDW) interactions and has been successfully applied to study the deformation and buckling of carbon nanotubes (Cao and Chen 2011; Chen et al. 2006; Pantano et al. 2003), as well as gating transition of the *E. coli* and *Mycobacterium tuberculosis* MscL (Bavi et al. 2016b; Chen et al. 2008; Tang et al. 2006). Figure 3a shows examples of the interaction energy between different protein components in closed and open crystal structures and the fitted curves used in FEM. The consistency between the interaction energy curves of the closed and open states suggests that the fitted FEM parameters are fairly transferable. Table 1 summarizes the fitted values of C , d_0 , and the exponents (m , n) for the closed crystal structure. When calculating

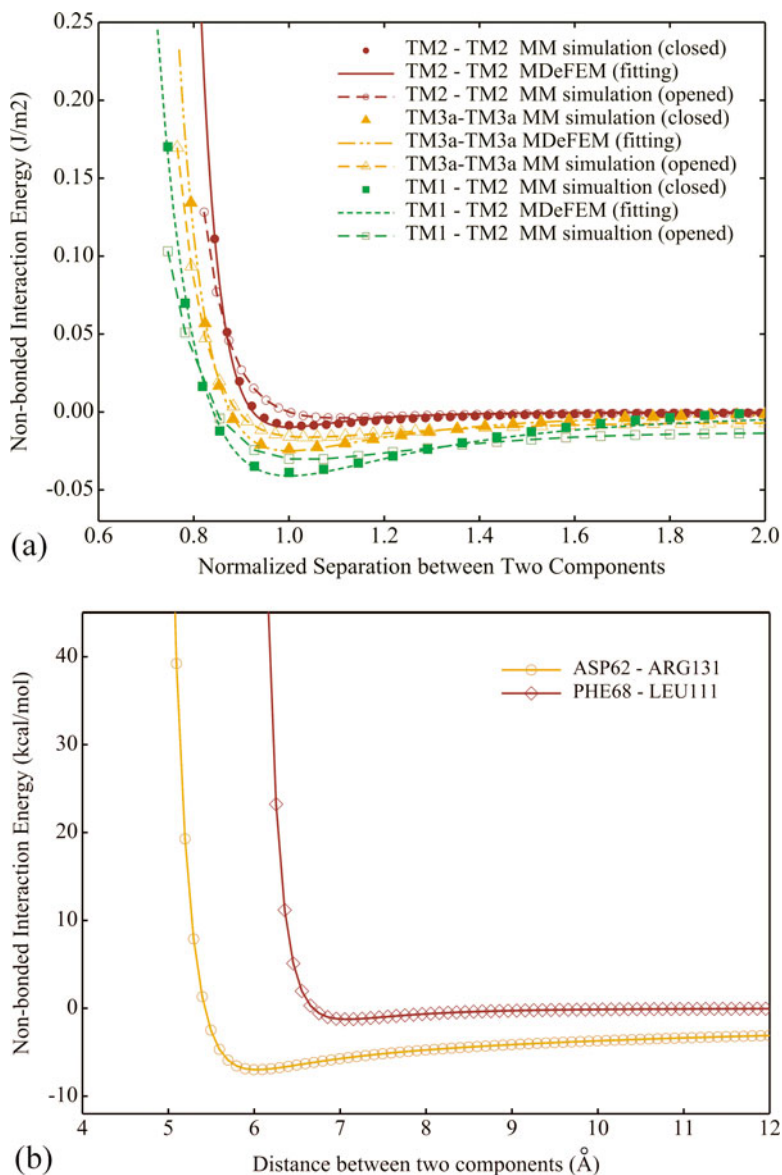


Fig. 3 Examples of the fitting of nonbonded interactions between helices of *E. coli*-MscS (a) and the interaction energy for two key interaction pairs (b). In (a), the distance between two components is normalized by their equilibrium distances. Data points are obtained through molecular mechanics (MM)

the interaction between TM2 and TM3a, the Phe68-Leu111 pair is excluded since it is already considered separately as discussed above. The constitutive interaction relationship is called in each increment through the analysis based on the relative position between two surfaces that are continuously updated. Similar

Table 1 FEM fitting parameters for the nonbonded interactions

Interaction pairs	d_0 (Å)	$\psi = 6Cn/d_0$ (GPa)	m	n
Lipid-TM1	5.5	4.0	8	3
Lipid-TM2	5.5	3.0	6	3
TM1-TM1	11.0	0.35	24	8
TM2-TM2	7.7	1.05	19	8
TM3a-TM3a	3.6	1.5	12	4
TM3b-TM3b	6.6	1.59	6	5
TM1-TM2	4.0	2.25	18	3
TM1-TM3a	10	29.8	24	23
TM2-TM3a	4.5	7.46	12	8
TM2-TM3b	7.7	7.48	24	20
TM3a-TM3b	3.0	2.6	11	2

to the connector elements, the pressure-based nonlinear interaction is defined both within the same monomer and across neighbor monomers as long as the distance between two components is smaller than 16 Å.

The different treatment of the key interaction pairs (described by node-to-node connectors) and other “less important” interactions (described by the averaged surface-to-surface pressure) allows a simple description of the continuum model while providing the opportunity to explore how these key interactions affect gating. A typical tension simulation of the gating transition of MscS takes only about 1 h, on a Thinkpad laptop with four 2.5 GHz CPUs and 8 GB of RAM.

Results and Discussion

In this section, gating transition of MscS in response to membrane stretch (tension) is firstly obtained. The open FEM structure is compared to the crystal open structure, and the intermediate structure is identified in FEM. Similar analysis is then conducted with some key interaction pairs (Asp62-Arg131 or Phe68-Leu111) or loops excluded to explore their role in the gating transition. Next, the kink between TM3a and TM3b is considered as helical (thus having larger cross section and Young’s modulus) rather a loop, leading to the discussion of a plausible inactivation mechanism. Finally, the effects of transmembrane voltage are analyzed.

MscS Gating Pathway

Shown in Fig. 4 are several snapshots of MscS during the membrane tension-driven gating process in comparison to the crystal structures for the closed (Bass et al. 2002) and open states (Wang et al. 2008). The results indicate that, during gating, all helices shift radially away from the center of the pore in a manner reminiscent of a mechanical camera iris, similar to the gating transition described for MscL (Betanzos et al. 2002; Tang et al. 2008). The TM1 and TM2 helices in

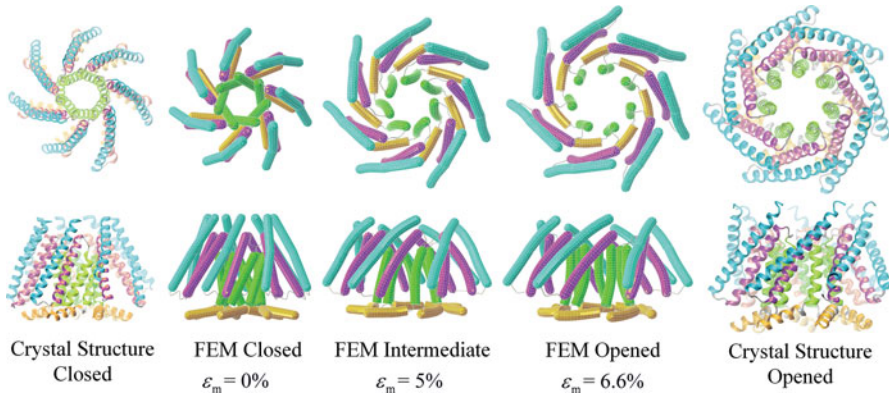


Fig. 4 Comparison between MscS gating pathways under equi-biaxial membrane tension between FEM results and the crystal models in both top (the top panel) and side views (the lower panel)

the same monomer move concurrently almost as a rigid body due to the strong inter-helical interactions. The TM3a helices rotate around and move outward from the central axis and finally become parallel to each other and normal to the membrane plane. Notably, the loop that connects TM1 and TM2 in the same monomer transitions from interacting with the end of TM3b in the neighboring chain (in the counterclockwise direction) to interact with the end of TM3b within the same monomer. Besides, significant rotation of TM3b helices around the central pore axis is observed (Fig. 4), though the average distance from TM3b ends to the pore axis is essentially unchanged. The conformational transitions discussed above are reversible in the FEM simulations once the membrane strain is removed.

On a qualitative level, the current FEM results are in agreement with the crystal structure models, regarding both the orientation and displacement of the transmembrane helices and connecting loops (Fig. 4). Small differences lie in the tilting of TM1 helices; from the top view (first panel in Fig. 4), the TM1 helices form a more compact bundle in the open crystal structure than in the FEM model. In addition to the open structure, intermediate structures during the gating process are obtained in the FEM simulation, and an example is given in Fig. 4.

In the FEM model of MscS in the closed form, the radius of the area lined by TM3a has a radius $R' \approx 7.0\text{\AA}$ (inset in Fig. 5), while the actual pore radius is $R \approx 2.4\text{\AA}$ (Wang et al. 2008) considering the pore's irregular inner surfaces. In the following, the pore radius (R) is estimated by $R = R' - 4.6\text{\AA}$ based on the FEM results of R' , and the pore diameter is thus $2R$. Figure 5 depicts the evolution of pore diameter during the gating process. As the membrane strain (ϵ_m) increases, the pore size of MscS increases slowly until ϵ_m reaches 5%, after which the pore diameter ($2R$) experiences a rapid increase leading to pore opening ($2R = 13\text{\AA}$) (Wang et al. 2008). Since the hydrophobic pore remains nonconducting until approximately $2R > 9\text{\AA}$ (the threshold diameter for hydration) (Beckstein et al.

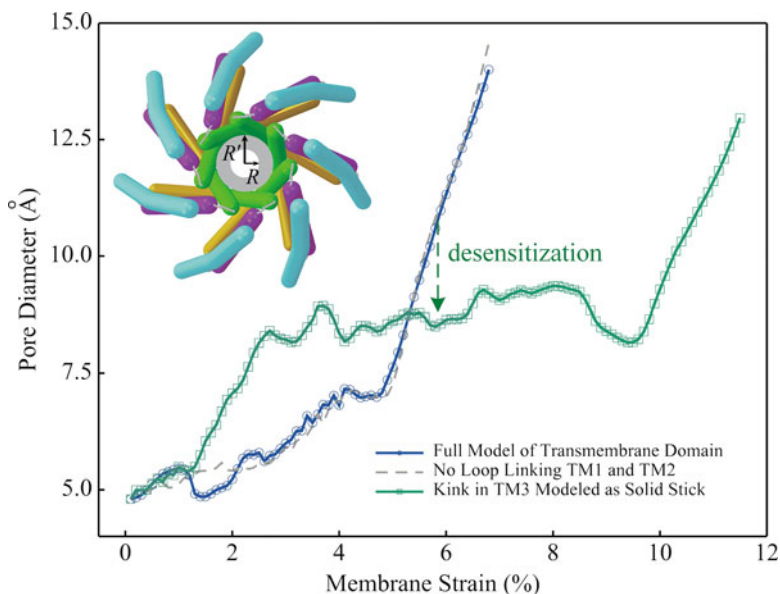


Fig. 5 Pore diameter evolution for different FEM models

2001), this observation is in qualitative agreement with the patch-clamp study of MscS (Martinac and Kloda 2003), in which the gating membrane tension for full opening of MscS was measured to be about half of that of MscL (Martinac and Kloda 2003). The gating membrane strain for MscL was estimated to be $\sim 10\%$ (Zhu et al. 2016), thus for MscS, the gating strain is expected to be 5%; the lipid membrane strain required for full opening of MscS in the present FEM simulation is 6.6% (Fig. 5), which is slightly larger than the 5% estimation.

Effect of the Key Interaction Pairs

Mutation studies have identified two key interaction pairs that greatly affect the gating behavior of MscS, and these involve the Asp62-Arg131 salt-bridge (Nomura et al. 2008) and the apolar interaction between Phe68-Leu111 (Belyy et al. 2010). For example, when the negatively charged Asp62 was replaced with either a neutral (Cys or Asn) or basic (Arg) amino acid, the gating threshold increased significantly (Nomura et al. 2008). Both F68S and L111S substitutions also led to severe loss-of-function phenotypes (Belyy et al. 2010). To provide a structural understanding of how these interactions influence the gating transition of MscS, we conduct FEM simulations with either pair of interactions excluded.

Firstly, the electrostatic interaction between Asp62-Arg131 in the continuum model (Fig. 1) is removed, while the van der Waals interaction between them is preserved (Fig. 6a) so as to mimic the replacement of Asp62 with a charge neutral amino acid. The structural transition upon membrane stretch up to 6.6% is shown in Fig. 6b–c. Without the strong electrostatic interaction, the TM1-TM2 helices tend to

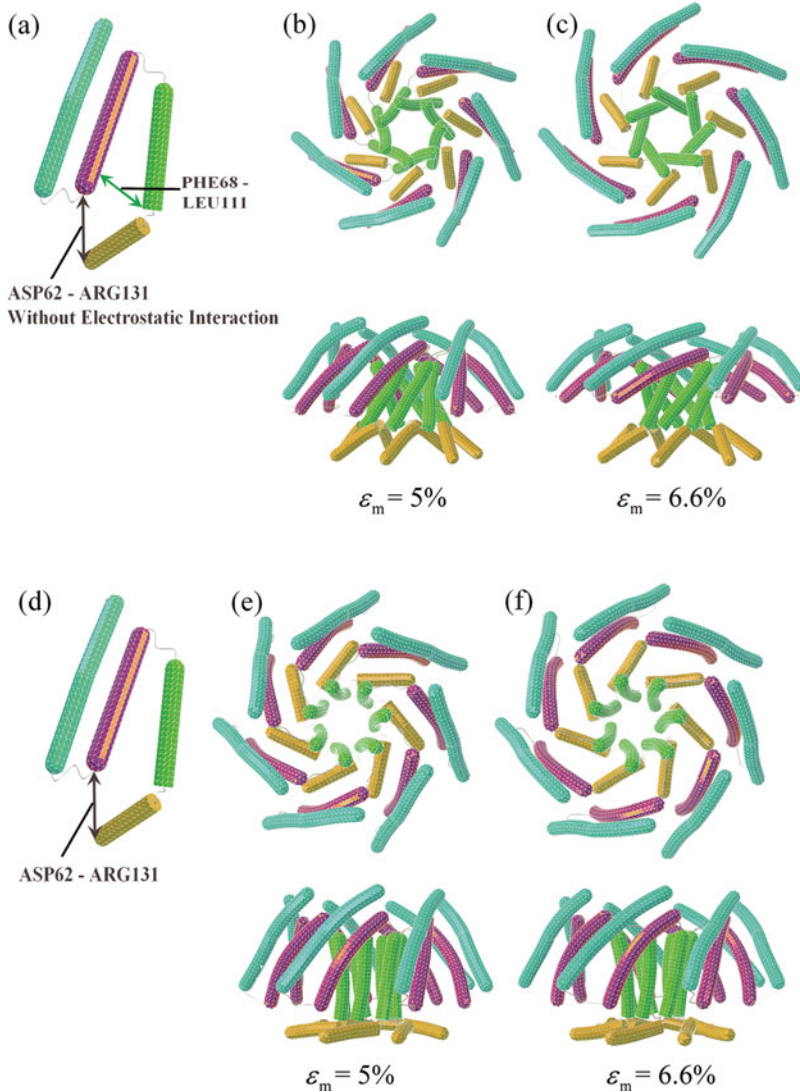


Fig. 6 FEM results of the structural transition of MscS under equi-biaxial membrane tension without electrostatic interaction between Asp62 and Arg131 (a–c) or without the Phe68-Leu111 traction (d–f)

detach from the end of TM3b. Under equi-biaxial membrane tension, all helices are lifted up and the kink in TM3 is straightened. As a result, the size of the channel pore surrounded by TM3a helices remains essentially unchanged despite the outward displacements of TM1 and TM2. Therefore, opening of the pore in this “mutant” requires an exceedingly large membrane strain (data not shown).

Similarly, a simulation without the apolar interaction between Phe68-Leu111 is conducted (Fig. 6d–f). During lipid membrane stretching, the TM1 and TM2 helices move outward, but the TM3a helices do not follow and the kink angle in TM3 remains unchanged. With a membrane strain of 6.6%, the pore size actually becomes smaller since the lower ends of TM3a helices are more closely packed because of the loss of TM2-to-TM3a attraction (F68-L111).

In short, the FEM simulations recapitulate the expected behaviors of the relevant mutants in which the key interactions are perturbed (Belyy et al. 2010; Nomura et al. 2008). The observed structural evolution in the FEM simulations also provides a physical understanding of how these interactions contribute to the gating transition. These results support the use of the FEM model to explore other contributions, to which we turn to next.

Effect of Structural Motifs (Loops) in Force Transmission

In this section, we explore the roles that periplasmic loops play in MscS gating by repeating FEM simulations with specific loops excluded from the model. When the loops connecting TM1 and TM2 helices are removed (with the Asp62-Arg131 interaction reserved), the structural response (data not shown) closely resembles that of the full channel model; likewise, the pore diameter evolution is not significantly perturbed (see Fig. 5). When the loops connecting TM2 and TM3 helices are removed, while the rearrangements of TM1 and TM2 helices are similar to those of the full channel model (Fig. 4), the TM3a helices collapse at the upper ends, leading to an essentially closed pore (Fig. 7). This suggests that the loops linking TM2 and TM3 helices are essential to the force transmission from lipid membrane to the pore-lining helices during MscS gating; this is quite different from the situation of MscL (Tang et al. 2008), in which loops that connect the transmembrane helices generally constrain channel opening (Ajouz et al. 2000; Tang et al. 2008). To the best of our knowledge, the importance of the loop between TM2 and TM3 helices to MscS gating has not been pointed in the literature, and this prediction can be further tested by studying loop deletion mutants in patch-clamp experiments as done for MscL (Ajouz et al. 2000; Bavi et al. 2016b).

Effect of the Helical Propensity of the TM3a-TM3b Kink: Inactivation

The Gly113 introduces a kink in TM3, and the region exhibits a weak helical propensity; thus, the kink segment of Gly113 is modeled as a loop (with an arc length of ~ 8 Å corresponding to the length of the backbone of Gly113) with a small cross-section and a lower Young's modulus (Fig. 1) in models described so far. To explore the significance of the helical kink, FEM simulations are performed in which the Gly113 region in TM3 is treated as helical but remained (enforced) as the bent shape and modeled as a curved solid stick (Fig. 8a) whose material properties take that of the TM3b helix. The structural response of this modified channel model under membrane tension is depicted in Fig. 8a and the pore diameter evolution process in Fig. 5 (square curve). When the membrane strain reaches 6.6%, the pore expands from 4.8 Å to 8 Å in diameter (Fig. 8a). At this point, the expanded pore still remains nonconducting since, for a hydrophobic pore, the threshold diameter

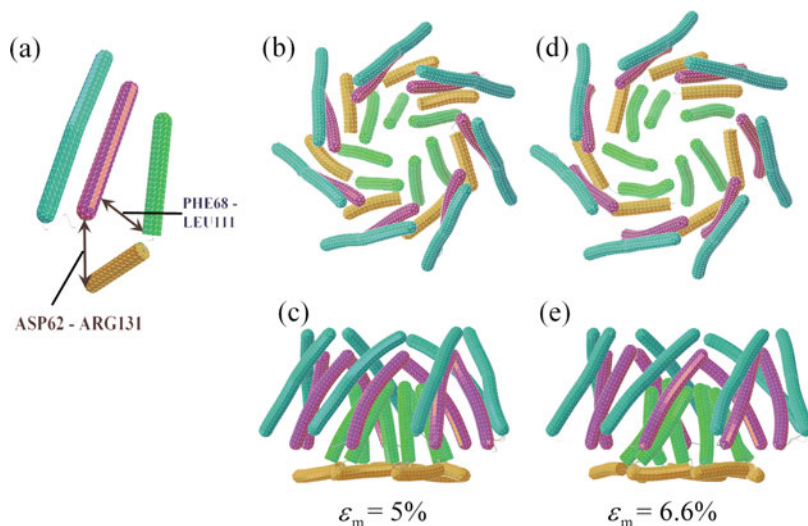


Fig. 7 Structural transition of MscS under equi-biaxial membrane tension with the TM2-TM3 loop removed

for hydration is $\sim 9 \text{ \AA}$ (Beckstein et al. 2001). To fully open the channel ($2R=13 \text{ \AA}$), a high membrane strain of 11.4% is required (Fig. 8a); this value is comparable to the gating strain of MscL (Zhu et al. 2016). Combining Fig. 4 and Fig. 8a, it seems that a flexible end of TM3a at the kink allows a more effective pulling by the Phe68-Leu111 interaction, thus leading to a lower gating strain. The interaction between TM3b and the cytoplasmic β domain where the TM3b- β interface is quite hydrophobic (Koprowski et al. 2011) promotes the kink around Gly113 (Koprowski et al. 2011; Petrov et al. 2013), which, in the closed state, may prevent backbone hydrogen bonding in the kink area, thus weakening the constraint at the lower ends of TM3a. The TM3b- β interface is indicated in Fig. 8a (ellipse), though noting that the cytoplasmic part is not explicitly included in the present study. Introducing higher polarity (charged residues) into the TM3b- β interface (e.g., the N117 K or G168D mutant) results in a weaker interaction between TM3b and the cytoplasmic β domain (Koprowski et al. 2011), which could be in favor of the backbone hydrogen bonding in the kink area. It has been observed that neither N117 K nor G168D mutant shows visible inactivation behavior, while both of them require a much larger membrane pressure for gating, which is consistent with the present results (Fig. 8a).

With these observations, it is tempting to speculate an inactivation mechanism for MscS: in the resting state, the kink of TM3 behaves like a loop, and upon rapid membrane tension, the channel is easily opened (Fig. 4 and the circle curve in Fig. 5). Under prolonged exposure to subthreshold membrane tension, however, the channel desensitizes (dashed arrow in Fig. 5) into a nonconductive state (the circle curve); this occurs because of the local transition of the kink region into a

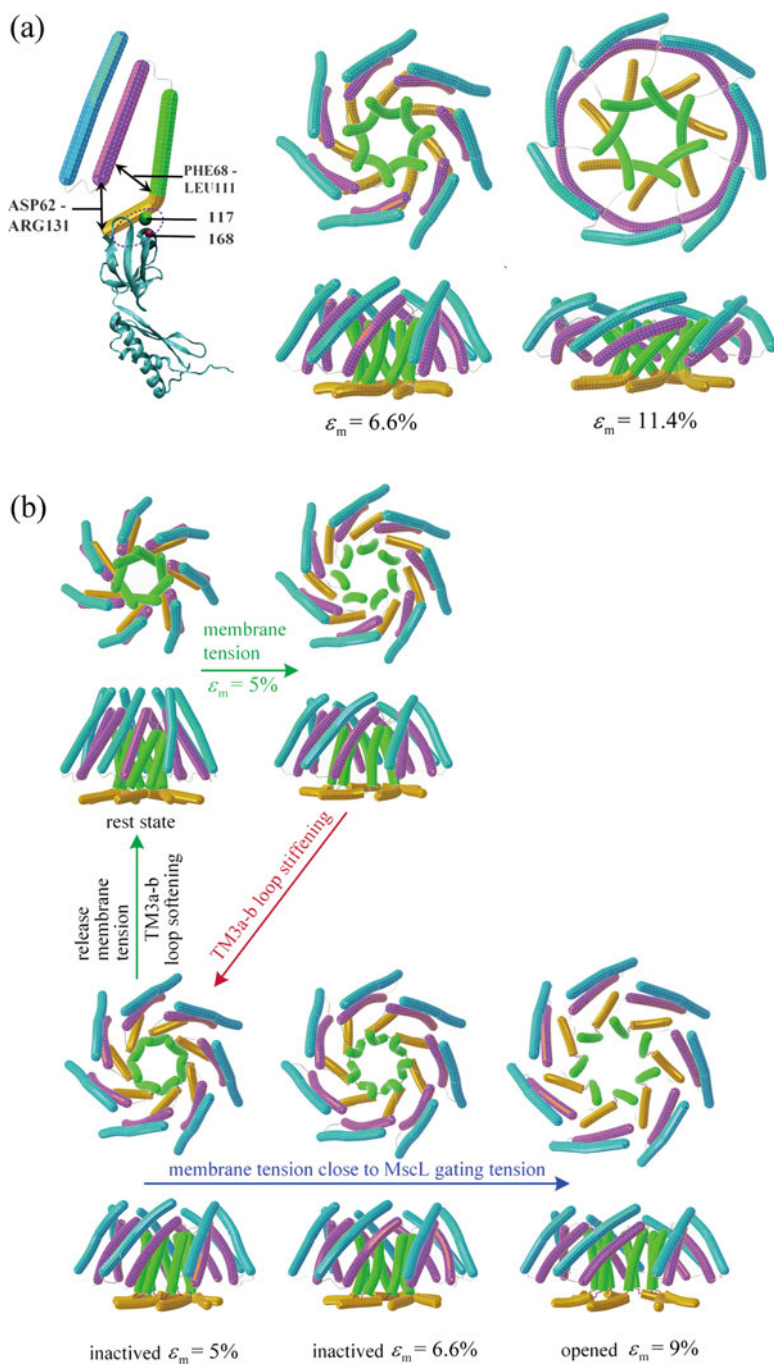


Fig. 8 (continued)

helical conformation that hinders further pore opening. The helical transition of the kink could be induced by the weakening of the TM3b- β interface because of the increased hydration around the interface domain caused by the rotation and outward movement of TM3 helices (Wang et al. 2008) or the reduction of lipids in the TM pockets (Pliotas et al. 2015).

To further test this hypothesis, a FEM simulation is conducted in which, at a subthreshold membrane strain, the Young's modulus of the TM3a-b loop is increased to the extent that its bending stiffness is equal to that of the TM3b helix. In Fig. 8b, after stretching the initial model (model in Fig. 5) by 5% ($\epsilon_m=5\%$), the TM3a-b loop is stiffened, which leads to reduced pore size; further opening the "inactivated structure" requires a very large membrane strain (9%) that is close to the gating strain of MscL. Increasing the stiffness of the TM3a-b loop at other reduced values ($<5\%$) of membrane strain leads to similar results. By contrast, if the stiffening of the TM3a-b loop occurs at a much larger membrane strain (e.g., $\epsilon_m=6.6\%$), the impact on the subsequent pore evolution is minimal. These results are in qualitative agreement with the experimental observation of Sukharev et al. (Kamaraju et al. 2011) that MscS inactivates primarily from the nonconducting state and channel opening prevents inactivation. It should be noted that the present quasi-static approach does not show the dependence of inactivation on (long) time (Kamaraju et al. 2011), which could be a result of desensitization against hydration inside the TM3a pore.

The mechanism for MscS inactivation has remained a mystery for decades. Though some key interactions and residues have been proposed to be involved, how these interactions and residues induce inactivation is poorly understood. The current study suggests that this behavior of MscS could be related to the conformational state of the helical kink at Gly113; a loop conformation allows other interactions to effectively pull TM3a helices outward thus opening the channel, while a more ordered helical conformation hinders the pulling, leading to channel inactivation. Previous experimental studies highlighted the importance of the Gly113 region as well. For example, another MscS-like channel from *Silicibacter pomeroyi* (MscSP) has a conserved glycine residue at the position equivalent to Gly113 in MscS (Petrov et al. 2013). However, the N117 residue in MscS is replaced by a charged residue Glu in MscSP based on the alignments of MscS and MscSP (Petrov et al. 2013), which leads to a weaker interaction between TM3b and the cytoplasmic β domain in MscSP (Koprowski et al. 2011; Petrov et al. 2013) and may facilitate the helical transition of the kink region, resulting in a higher gating threshold (square curve in Fig. 5); indeed, it was reported that the threshold gating strain of MscSP is ~ 1.5 times of that of MscS (Petrov et al. 2013). In another two MscS-type channels,



Fig. 8 Structural transition of MscS under equi-biaxial membrane tension with the TM3a-TM3b kink considered as helical (a) and the proposed mechanism for gating, desensitizing, and inactivation of MscS. The purple ellipse in (a) indicates the TM3b- β interface. The increased helical propensity of the kink could be induced by weakening the interaction between TM3b and the cytoplasmic β domain (e.g., the G168D mutant), thus, to an extent, relaxing the kink

MscMJ of *M. jannaschii* (Kloda and Martinac 2001) and MscCG of *C. glutamicum* (Borngen et al. 2010), inactivation was not observed (Petrov et al. 2013); this might be explained by the observation that Gly113 in MscS is replaced in the equivalent position by Asp and Ser in MscMJ and MscCG, respectively (Petrov et al. 2013).

In addition to the connection to these previous studies, the inactivation mechanism proposed here (Fig. 8b) can be tested by replacing Gly113 with amino acids of higher helical propensity or extending the kink area (e.g., multiple Gly insertion) to further increase its structural flexibility; the latter was done by Martinac et al. for MscL (Bavi et al. 2016b). Both mutations are expected to result in the absence of inactivation in MscS.

Interestingly, an earlier work (Akitake et al. 2007) has conducted a series of mutation studies both at the Gly113 kink region and the TM3b- β interface area (Gly121), though that study was aimed to show that the Gly113 kink is a unique feature of the inactivated state and the closed structure favors buckling at Gly121, which was not supported by later studies (Pliotas et al. 2012, 2015; Ward et al. 2014). Thus we here attempt to interpret the results of mutation studies in Akitake et al. (2007) from the perspective of our proposed inactivation mechanism. First, the G113A mutant in Akitake et al. (2007) had increased helical propensity in the kink region and indeed showed no inactivation behavior; this agrees with our inactivation mechanism. The gating tension of G113A mutant was, however, observed to be comparable to that of the wild-type MscS in Akitake et al. (2007); this observation is not consistent with the predicted gating pathway in Fig. 5 (square curve), probably because the Ala substitution interacts with phospholipids inside the TM pockets (Pliotas et al. 2015) (not considered in the present model), thus strengthens the apolar interaction between the lower end of TM2 and that of TM3a (the main traction that pulls TM3a bundle open (Fig. 6)); alternatively, the Ala substitution may lead to structural rearrangements not considered in the present model (e.g., a certain extent of kink straightening (Akitake et al. 2007)) that enhance the interaction between the ends of TM2 and TM3a. To test these possibilities, a simulation is conducted where the kink region is treated as helical and modeled as a curved solid stick (Fig. 8a), while at the same time the TM2-TM3a traction force (green arrow in Fig. 8a) is increased by 50%. The gating strain is found to be 6.0% in this case, close to the gating strain (6.6%) of the model in Fig. 4 where the kink is modeled as a loop.

Second, the Q112G mutation (Akitake et al. 2007) increased the flexibility of the kink area, but instead of resulting in the absence of inactivation as suggested in our proposed inactivation mechanism, the Q112G mutant exhibited faster inactivation (Edwards et al. 2008). We conjecture that the Q112G mutation may weaken the apolar interaction between the lower end of TM2 and that of TM3a and may make the kink area “too flexible” that upon channel opening or even in the initial state, water molecules enter the TM pocket between TM1-TM2 and TM3a thus further causing the detachment between TM2 and TM3a. To test this assumption, a simulation is conducted for a modified Q112G mutant model, where the kink region is treated as a loop and modeled as a flexible quasi-one-dimensional winding beam (Fig. 1b), while at the same time the TM2-TM3a traction force (green arrow

in Fig. 1b) is decreased by 10%. The simulation results in a structure close to that in Fig. 6 where the Phe68-Leu111 traction is missing and the channel cannot be opened, which may explain the faster inactivation of Q112G. As replacing one residue could change a set of interactions nearby (Wang et al. 2008), especially for the subtle situation at the Gly113 area which also affects the TM2-TM3a apolar interaction or possibly the hydrophobic lock of the channel pore (Anishkin and Sukharev 2004; Vora et al. 2006), mutation studies in this region require careful analysis and interpretation. The two simulations mentioned above serve as exploratory models by controlling the state of the kink (helix vs. loop) and the traction between the ends of TM2 and TM3a. More detailed experimental and simulation works are required in future to determine the structural and interaction changes around the kink area due to mutations.

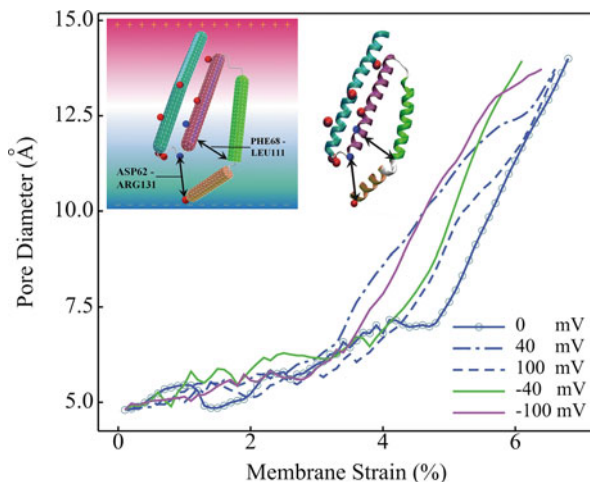
At the TM3b- β interface, two mutations were conducted in Akitake et al. (2007), G121A and A120G. The G121A substitution, as discussed above, can reinforce the hydrophobic TM3b- β interface, promoting the kink around Gly113 (Koprowski et al. 2011; Petrov et al. 2013) and resulting in a more flexible kink region at Gly113. As expected from our mechanism, the G121A mutant was observed to open easily and showed no inactivation (Akitake et al. 2007). The A120G mutation, on the other hand, weakened TM3b- β interface, thus promoting the helical propensity at the kink area. Again as expected from our mechanism, the A120G mutant exhibited a high degree of inactivation (Akitake et al. 2007). Therefore, the results for both G121A and A120G mutants are supportive of the inactivation mechanism proposed here.

Effect of Transmembrane Voltage

Transmembrane voltage is applied to the continuum model to investigate its effects on MscS gating (Akitake et al. 2005; Nomura et al. 2016). For simplicity, we only consider the charged residues in the atomic model which are mapped to the finite element nodes in the continuum model (inset of Fig. 9). There are only a few charged residues in the transmembrane domain, which are Arg46, Arg54, Arg59, Lys60, Asp62, Asp67, Arg74, Arg88, and Arg128. Transmembrane voltages from -100 mV to $+100$ mV (Akitake et al. 2005) are applied with the exterior of the cell being positive with respect to its interior. No evident dependence of the activation pathways (Fig. 9) on transmembrane voltages is observed in the FEM simulations. This finding is consistent with recent experimental results (Akitake et al. 2005; Nomura et al. 2016). The rectifying behavior that the conductance is larger at positive voltages observed in some studies (Martinac et al. 1987; Petrov et al. 2013) could be a result of the inside-out nature in patch-clamp experiments and the slight anion preference of MscS (Martinac et al. 1987; Sukharev 2002).

Another observation of depolarizing voltage-dependent inactivation (Akitake et al. 2005), however, remains unsettled. How negative voltages promote inactivation rate must await further studies with more sophisticated models combined with experimental analyses. It is possible that, though the electric field causes only small displacements on the charges in MscS during activation (Akitake et al. 2005), its presence may disrupt some key interactions like the TM3b- β interface (e.g., Asn117). Besides, voltage-dependent inactivation may be related to the

Fig. 9 Pore diameter evolutions of MscS under equi-biaxial membrane tension and different transmembrane voltages. Residue charges and electric field are incorporated into the continuum model based on the charge distribution in the crystal structure (inset). Red particles represent positive residues and blue particles negative residues



conformational changes of the cytoplasmic domain (not included in the present study) as indicated in Rowe et al.'s work (Rowe et al. 2014) where the G168D mutant is insensitive to applied voltages.

Coupled Continuum Mechanical-Continuum Solvation Approach with Application to Gating Mechanism of MscL

Models and Methods

In this section, the construction process of the continuum mechanics model for *E. coli*-MscL is presented in detail. Relevant material and interaction parameterizations are either calibrated by matching results at the atomistic and continuum level or obtained from a previous study (Chen et al. 2008). The computational framework includes essentially two components: continuum mechanics (CM) simulations based on finite element method (FEM), which solve for the deformed structure under specific external loads, and the continuum solvation model (CS), which computes solvation forces based on structural information from the CM model. The commercial software ABAQUS (2011) (for CM calculations) and open-source software APBS (Baker et al. 2001) (for CS model) are used in this fully integrated (CM/CS) simulation framework. Detailed ABAQUS-APBS co-simulation protocols are presented.

Continuum Modeling of *E. coli*-MscL and Interactions Within the Protein

Although *E. coli*-MscL is one of the most studied MS channels, the only available X-ray crystal structure in the literature is for the MscL from bacteria *Mycobacterium tuberculosis* (*Tb*), which was captured in its closed state by Rees Lab (Chang et al. 1998). By retaining the main features of the crystal structure of *Tb*-MscL, the atomic

structure of *E. coli*-MscL was developed based on homology modeling along with other experimental constraints (Sukharev and Anishkin 2004; Sukharev et al. 2001a; Sukharev and Corey 2004); the closed state model is shown in Fig. 10g. The crystal structure of *Tb*-MscL was further refined by Stefan Steinbacher et al. (2007), where its N-terminal (termed S1) was an amphipathic helix positioned approximately parallel to the cytoplasmic surface of the membrane. This refined structure for *Tb*-MscL was supported by later studies (Iscla et al. 2008). Based on the new structure for *Tb*-MscL and MD simulations, S1 domain was suggested to interact closely with the lipid membrane, which may facilitate the opening of the channel (Iscla and Blount 2012; Vanegas and Arroyo 2014). While developing the new atomic *E. coli*-MscL structure based on the revived *Tb*-MscL crystal structure (Steinbacher et al. 2007) is beyond the scope of the present study, it should be noted that the effect of S1 domain as a “sliding anchor” to the lipid membrane thus helping channel opening (Iscla and Blount 2012; Vanegas and Arroyo 2014) is not available for studying in this work since we are using Sukharev S. et al.’s model (Sukharev and Anishkin 2004) whose S1 domain assembles as a helix bundle. The S1 domain’s interaction to the lipid and its hydration process may be to some extent different from what we present in the following context but is not expected to bias the principal outcomes of the current model.

An *E. coli*-MscL molecule is formed by five chains (from chain-1 to chain-5) assembled as a fivefold structure around its symmetry axis, and each single chain (Fig. 10a) of MscL consists of four helices (referred as TM1, TM2, S1, and S3) and several loops. Within the transmembrane helix bundles, the five TM2 helices form the outer boundary that interact extensively with the lipid membrane, while the longer TM1 helices form the inner boundary for the pore and have limited contact with the lipid. The TM1 and TM2 helix bundles share the same fivefold symmetry axis, denoted as the z -axis here, which is also the direction of the membrane normal. Each TM1 helix and its nearest neighbor, a TM2 helix, are assembled into a pair through a periplasmic loop. A closer view of the TM1 shows that there is a break near the top of the helix due to Pro-43; thus, the segment above Pro-43 is also referred to as the S2 helix. The cytoplasmic region contains two different types of helix bundles, referred to as the S1 helices and S3 helices; each bundle contains five subunits, and each subunit of the S1 or S3 bundles is connected, respectively, to a TM1 or TM2 helix through a loop. The TM1, TM2, S1, and S3 helices correspond to residues Asn-15–Gly-50, Val-77–Glu-107, Ile-3–Met-12, and Lys-117–Arg-135, respectively (Sukharev and Anishkin 2004). In the subsequent discussion, unless otherwise specified, MscL refers to the *E. coli*-MscL.

In the continuum model, we start from the backbone structure with side chains removed because when side chains are presented, some subcomponents are unable to be separated, such as the TM1 and TM2 helices in the same chain which are very close to each other. More dedicated calculation of the surfaces of the protein could be considered in future work to get a more precise model. The molecular surface (Fig. 10c) of each single chain of MscL molecule is calculated and triangularized using the MSMS program (Sanner et al. 1996) with appropriate probe radius to avoid overlapping of subcomponents (such as TM1 and TM2 helices). Among the multiple subcomponents of MscL, helices are modeled to have larger cross sections because

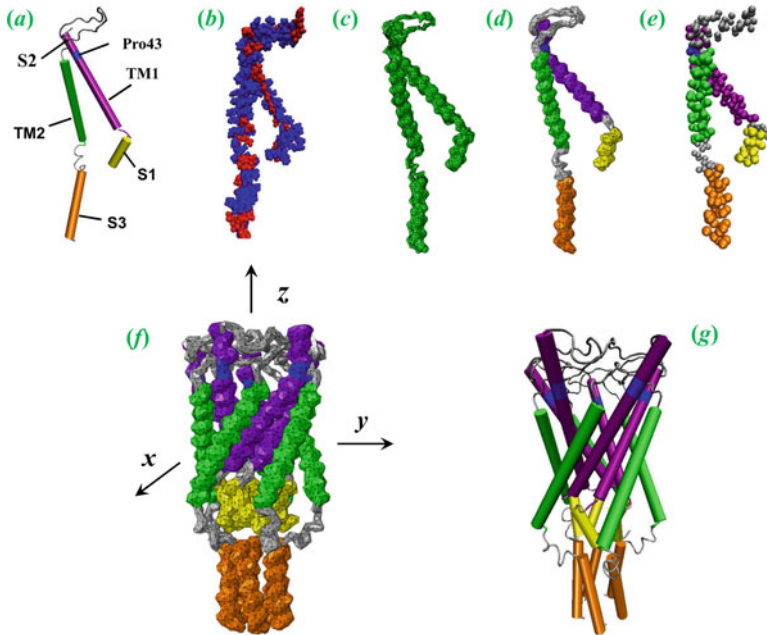


Fig. 10 Illustration of various steps in constructing the FEM Model of *E. coli*-MscL. (a, b) the atomic structure of one chain of *E. coli*-MscL in cartoon and van der Waals (VDW) representation where blue indicates hydrophobic and red hydrophilic; (c) the triangularized molecular surface; (d) the simplified triangularized molecular surface and the volume enclosed by this surface are discretized into tetrahedral elements; (e) the coarse-grained model of one chain of *E. coli*-MscL; (f) the FEM model of *E. coli*-MscL; (g) structural model of *E. coli*-MscL based on homology modeling (Sukharev and Anishkin 2004). Zhu et al. (2016), reprinted with permission of Springer

they are strengthened by the hydrogen bonding between residues, while the loops are modeled much thinner (Fig. 10c). The triangularized surface is then simplified by the QSLIM program (Heckbert and Garland 1999) to reduce the number of the surface triangles to 2000 (Fig. 10d). The volume enclosed by this simplified surface is subsequently discretized into a 3-D mesh consisting of tetrahedral elements. The final FEM model consists of 2,685 nodes and 9,699 finite elements for each chain of MscL.

Assuming that the mechanical properties of each component vary little with respect to sequence, the properties of each component are assumed to be homogeneous, isotropic, and constant during the gating process. The only exception is for the break between S2 and TM1 (Pro43, illustrated by the blue segment in Fig. 10a), whose properties are determined separately and less stiff than the rest of the helix. The material properties of each component are calibrated by matching results of normal mode analysis (NMA) at the continuum and atomistic levels. The NMA for individual components of the channel is conducted in vacuum using the Gromacs MD simulation package with the Gromos96 vacuum parameter set (Van Der Spoel et al. 2005) so as to be consistent with the continuum calculations which are also

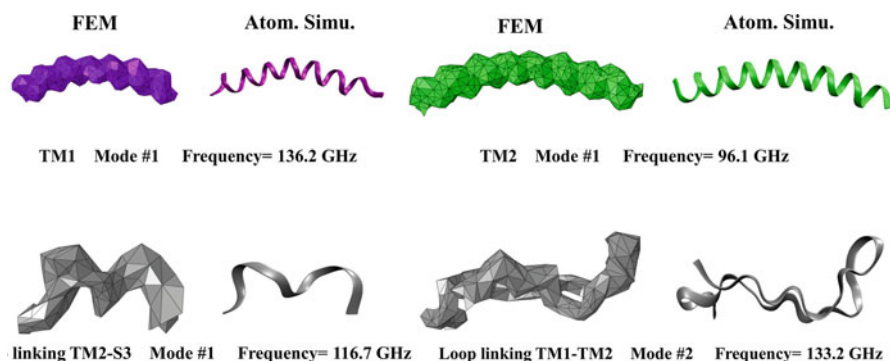


Fig. 11 Examples of several lowest eigenmodes and frequencies of helices and loops. Results of molecular mechanics are compared to those of the finite element simulations. Here the TM1 helix only includes to the segment below Pro43. Zhu et al. (2016), reprinted with permission of Springer

Table 2 Phenomenological material properties of the continuum components

Properties of MscL	Helices						Loops			Lipid membrane	
	TM1	TM2	Pro43	S1	S2	S3	TM1-TM2	TM1-S1	TM2-S3	Head groups	Tails
Young's modulus E (Gpa)	69	54	11	9	14	14	20	10	12	0.124	0.064
Poisson's ratio ν	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.5	0.5

conducted in vacuum. A possible way to get more realistic elastic properties is to conduct NMA considering the water and lipid environment in both continuum and atomistic calculations though this kind of NMA at the continuum level is highly complex. The Young's modulus is then varied at the continuum calculations such that the eigenvalues and eigenvectors for the three lowest-frequency modes best fit the results from the atomistic normal mode calculations. The lowest eigenmode of the TM1 or TM2 helices, for example, is essentially flexural bending as shown with both continuum and atomistic configurations in Fig. 11. The lowest frequencies are 136.2 and 96.1 GHz for TM1 and TM2 helices, respectively, which lead to the fitting of their effective Young's moduli as 69 GPa (TM1) and 54 GPa (TM2). Summarized in Table 2 are the key mechanical properties for the helices and loops. These are smaller than estimated in the previous work (Chen et al. 2008) where the helices, for instance, were modeled as thinner (with a diameter of 5 Å) elastic cylinders by considering only the main chain and larger than the estimated range of the Young's modulus of MscL α -helices (0.2 to 12.5 GPa) in Martinac et al.'s recent work (Bavi et al. 2017) with constant-force steered molecular dynamics (SMD).

The nonbonded interactions within MscL are treated by adopting a sufficiently detailed yet computationally efficient protocol to replace the oversimplified surface-to-surface contact model used previously (Chen et al. 2008; Tang et al. 2006). To reduce computational cost, the new approach first simplifies the full-atomic

structure into a coarse-grained (CG) model consisting of a much smaller number of particles based on the Martini force field v2.1 (Monticelli et al. 2008). This CG process reduces the number of particles from 2166 to 283 for each chain of *E. coli*-MscL with side chains included (Fig. 10e). The coarse-grained particles are then mapped to the nodes of the continuum MscL model so that it consists of two types of nodes: ordinary finite element (FE) nodes and the “chemical nodes.” The “chemical nodes” in the continuum MscL model refer to the FE nodes that have the same coordinates to that of the CG particles (Fig. 10d, e). Interactions between the Martini particles are then applied to the corresponding FE nodes in the continuum mechanics simulation. These interactions include the particle-to-particle interactions between different components within the same chain and between components in different chains. For example, within chain-1, one particle on the TM1 may interact with another particle on S3, TM2, S1, or loops of chain-1, and as for different chains, one particle on the TM1 of chain-1 may interact with any particle on chain-2 (from S3 to S1). In the MDeFEM framework, the interactions among atoms within each continuum component (e.g., particles on TM1 of chain-1) are not computed explicitly because the corresponding energy is already included in the elastic representation of the continuum components; this is one reason that the computational cost associated with the continuum framework is substantially lower than all-atom simulations. The nonbonded interactions between CG particles in different continuum components are calculated using pair-wise terms following the standard cutoff schemes commonly used in atomistic simulations. Specifically, we adopt a group of “connector elements” in the mechanical space, each of which characterizes the nonbonded interaction between a pair of interacting CG particles. These connector elements are not actually FEM elements but a special kind of invisible connectors in ABAQUS used to define force-based nonlinear interactions, just like the nonbonded interactions in MD. The connector element behavior is nonlinear elastic including two possible forms:

$$\begin{aligned} \text{Lennard-Jones form : } V_{LJ} &= 4\epsilon \left[(\sigma/l)^{12} - (\sigma/l)^6 \right] \\ \text{Coulomb's form : } V_{Elec} &= \frac{1}{4\pi\epsilon_0\epsilon_r} \frac{q_1q_2}{l^2} \end{aligned} \quad (2)$$

where ϵ and σ are the energy and distance parameters, respectively, to characterize the Lennard-Jones interactions. ϵ_0 is the vacuum dielectric constant, $\epsilon_r = 15$ is the relative dielectric constant for electrostatic screening in the Martini force field v2.1 (Monticelli et al. 2008), and q_i is the partial charge of the i -th particle. l denotes the distance between a pair of interacting particles. We note that with the polarizable MARTINI water model (Yesylevskyy et al. 2010), the relative dielectric constant is substantially smaller (2.5); test calculations are also done with this value, and the results are generally very similar (data not shown) since the behavior of MscL appears to be largely dictated by the nonpolar solvation (see discussions below). For the interactions among the MscL components, all the parameters involved in the Lennard-Jones interaction are obtained from the Martini force field v2.1; all distances are calculated based on the CG particle coordinates that are continuously updated and mapped from the continuum model. Also continuously updated are

the connector elements to include newly emerged interactions when two remote particles come closer during the simulation process. The Lennard-Jones interaction cutoff length is set to be 1.0 nm and the electrostatic cutoff length 1.2 nm.

The major advantage of the irregular molecular surface FEM model over the highly simplified cylindrical stick model (Chen et al. 2008; Tang et al. 2006) is that they take into account the molecular nature of MscL, such as its irregular shape. In addition, the mapping between CG particles and corresponding FEM nodes makes it straightforward to define “chemical nodes,” which encodes the key chemical characteristics of the molecule (i.e., charge distribution and solute/solvent interface) that are required in the solvation calculations.

Continuum Modeling of Lipid Bilayers and Interactions Between MscL and Lipid

Motivated by the natural difference in the chemical and physical properties of these regions, the lipid membrane bilayer is modeled as a sandwich panel that consists of three layers (Fig. 12): a soft layer in the middle with a thickness of 2.5 nm and two hard layers in two sides with a thickness of 0.5 nm each (Chen et al. 2008). These values for thickness estimation are based on density map of water and lipids from a MD simulation of the POPE lipid system (Gullingsrud and Schulten 2004), and both the derived head group thickness and the tail group thickness are consistent with general thickness estimations for the POPE lipid membrane. Meanwhile, the work of Andrew M. Powl et al. identified the hydrophobic thickness of *E. coli*-MscL associated with the lipid membrane as 25 Å (Powl et al. 2005b). Thus, there is a good match between the hydrophobic thickness of helix and bilayer; and hydrophobic mismatch is not considered in the present model.

For the case of in-plane membrane stretching, which is the major driving force for MscL gating (Tang et al. 2006), a flat square membrane (within the x - y plane) with a size of 400×400 Å is employed. The equi-biaxial membrane tension is most likely induced by osmotic pressure: assuming the liposome is spherical with a typical diameter of 1.0 μm , a patch of membrane with the size of 400×400 Å corresponds to a center angle of ~ 0.08 , which suggests that the curvature of the patch is negligible. To embed the channel into the continuum membrane, a cavity (hole) with the shape of a 10-petal flower (Fig. 12) is created in the middle of the membrane with the size and shape of the cavity conform to those of MscL transmembrane helices in the closed state with an equilibrium distance of ~ 5.5 Å, as was measured from the trajectories in previous all-atom simulation (Gullingsrud and Schulten 2003).

The three-layer phenomenological model (with a thickness of ~ 35 Å (Gullingsrud and Schulten 2003)) for the palmitoyloleoylphosphatidylethanolamine (POPE) lipid bilayer is first assumed to be homogeneous and isotropic, and its mechanical properties are obtained from a previous study (Chen et al. 2008) where the effective elastic properties of the head group layers and the tail group layer were derived from MD simulations (Gullingsrud and Schulten 2004). And these parameters are overall consistent with general simulation or experimental estimations (Binder and Gawrisch 2001; Chacon et al. 2015; Venable et al. 2015), though they may be further optimized through recent developments about the new

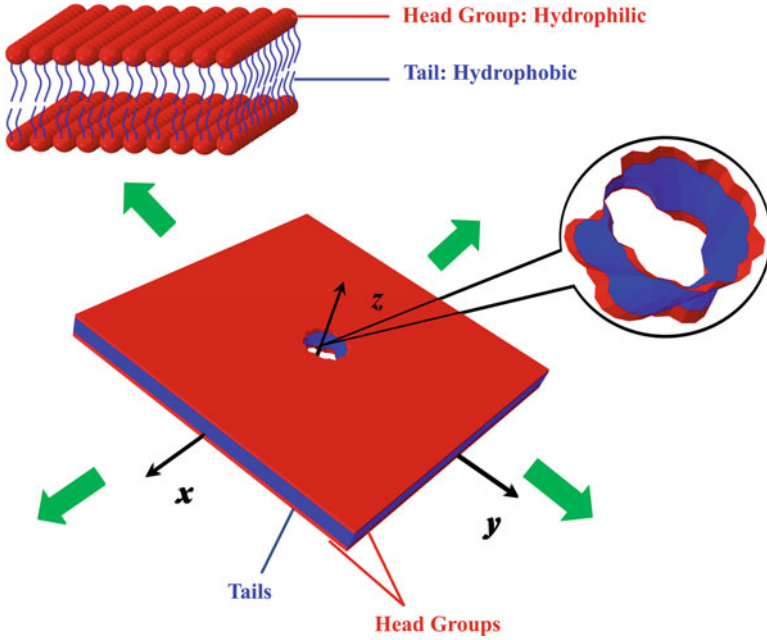


Fig. 12 Schematic of equi-biaxial tension of the lipid membrane model. A zoomed-in view is shown to illustrate the 10-petal lipid hole that encompasses the MscL protein in the CM model. Zhu et al. (2016), reprinted with permission of Springer

force decomposition methods (Torres-Sánchez et al. 2015; Vanegas et al. 2014). For isotropic linear elasticity, the constitutive relation (stress-strain relationship) is given by

$$\begin{Bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \\ \sigma_{23} \\ \sigma_{13} \end{Bmatrix} = \begin{pmatrix} \frac{E(1-\nu)}{(1+\nu)(1-2\nu)} & \frac{E\nu}{(1+\nu)(1-2\nu)} & \frac{E\nu}{(1+\nu)(1-2\nu)} & 0 & 0 & 0 \\ \frac{E\nu}{(1+\nu)(1-2\nu)} & \frac{E(1-\nu)}{(1+\nu)(1-2\nu)} & \frac{E\nu}{(1+\nu)(1-2\nu)} & 0 & 0 & 0 \\ \frac{E\nu}{(1+\nu)(1-2\nu)} & \frac{E\nu}{(1+\nu)(1-2\nu)} & \frac{E(1-\nu)}{(1+\nu)(1-2\nu)} & 0 & 0 & 0 \\ & & & G_{12} & 0 & 0 \\ & \text{sym} & & & G_{23} & 0 \\ & & & & & G_{13} \end{pmatrix} \begin{Bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ \varepsilon_{12} \\ \varepsilon_{23} \\ \varepsilon_{13} \end{Bmatrix} \quad (3)$$

where σ_{ij} and ε_{ij} are the stress and strain tensor components. For isotropic linear elastic materials, the shear moduli $G_{12} = G_{23} = G_{13} = \frac{E}{2(1+\nu)}$; thus, the elastic properties for each layer of the continuum lipid bilayer model are governed by Young’s modulus, E , and Poisson’s ratio, ν , listed in Table 2. To take into account the prominent characteristics of the fluidic lipid membrane, i.e., its incapability

of sustaining a large shear stress, we model each layer of the lipid membrane as orthotropic and reduce the in-plane shear modulus of the continuum lipid model, G_{12} (Eq. 3), to a small value (without losing generality, $\frac{E}{2(1+\nu)}/1000$ is adopted), while all other parameters are inherited from the previous isotropic continuum lipid model. Such a change does not affect the area compressibility under equi-biaxial tension, which is used to estimate the lipid membrane's Young's modulus and Poisson's ratio (Chen et al. 2008) because of the decoupling between tension and shear. It is expected that the in-plane fluidity of the lipid bilayer allows a closer interaction between the surface of the 10-petal lipid hole and the MscL molecule, thus facilitating the gating process since the in-plane lipid tension has been regarded as the major driving force that pulls the channel open (Moe and Blount 2005; Tang et al. 2006). Upon applied tension, the behavior of a real membrane may not be strictly elastic or orthotropic as described, and it was found in a molecular dynamics study that MscL inclusion may increase the rigidity of the membrane (Jeon and Voth 2008). For simplicity, these details are ignored in the present work and are not expected to have a significant influence on the structure or behavior of MscL based on our previous studies (Chen et al. 2008; Tang et al. 2008).

The nonbonded interactions between the lipid hole and the MscL molecule are treated as in Chen et al. (2008) where the interactions between helix and lipid are represented by a pair-wise effective pressure-distance relationship in the following form:

$$p(\alpha_i) = \frac{Dn}{d_0} \left[(d_0/\alpha_i)^{m+1} - (d_0/\alpha_i)^{n+1} \right] \quad (4)$$

where p and d_0 are the interaction pressure and equilibrium distance between two surfaces, respectively, and α_i is the instantaneous distance between two deformed surfaces for the i -th element. This nonbonded interaction model implicitly includes both electrostatic and van der Waals interactions and has been successfully applied to study the radial elastic properties of multiwalled carbon nanotubes (Chen et al. 2006) and the deformation and buckling of double-walled carbon nanotubes (Pantano et al. 2003) as well as nanoindentation of nanotubes (Cao and Chen 2006). To estimate the helix-lipid interactions, the insertion energy profile of a single helix (TM1, TM2, or S1) is calculated with an implicit membrane model; i.e., the helix is gradually transferred from the implicit membrane to the implicit bulk solution. Determined from molecular mechanics calculations (Chen et al. 2008), shown in Table 3 are the well-depth D and the exponents (m , n) which are fairly transferable to the current model since we are using the same atomic structure of *E. coli*-MscL.

It is noted the irregular surface of the lipid membrane is not included in this work, and the particle-to-particle lipid-channel interaction force is averaged to the

Table 3 Parameters for the nonbonded interactions between different helix-lipid pairs (Chen et al. 2008)

Interaction pairs	$d_0(\text{\AA})$	$\psi = 6Dn/d_0(\text{GPa})$	m	n
Lipid-TM1	5.5	2.0	9	3
Lipid-TM2	5.5	2.0	7	3
Lipid-S1	7.0	0.025	4	1

surfaces of them; thus, the heterogeneous binding of the lipid to the channel is lost. Further refinements of the continuum mechanics lipid model and its heterogeneous interaction to the channel could be considered in future work with different levels of sophistication.

Continuum Solvation Modeling

As described above, the new continuum MscL model is meshed in a way so that it consists of two types of nodes: ordinary finite element (FE) nodes and the “chemical nodes” which are also FE nodes but correspond to the CG particles. While all FE nodes contribute to the mechanical deformation of MscL, the “chemical nodes” are subjected to additional force contributions from solvation and inter-component interactions. Since the new continuum representation adopted here retains the information about the irregular shape of the protein and spatial distribution of the amino acid residues, the solvation contribution (electrostatic plus apolar solvation forces) can be readily calculated using a popular continuum solvation model. Shown in Fig. 10e are the CG particles of *E. coli*-MscL that bear solvation forces in the continuum solvation calculations. Following the mapping of these CG particles to the nodes of the continuum MscL model, the solvation forces are also transferred from the CG particles to the corresponding “chemical nodes” and included in the subsequent continuum mechanics simulations.

The total solvation free energy is usually decomposed into apolar and electrostatic contributions. For the electrostatic component, the nonlinear Poisson-Boltzmann (NLPB) (Davis and Mccammon 1990; Honig and Nicholls 1995) model is used. The van der Waals radius ($2^{1/6}\sigma/2$) of each CG particle is used to approximately determine the dielectric boundary between MscL and the solvent, and ions in solution have a finite radius of 2.0 Å. The spline-based (Im et al. 1998) molecular surface, which permits stable solvation force calculations, is used with a 0.3 Å spline window. The channel is surrounded by the nonpolar membrane environment, which is represented crudely by a low dielectric slab following the procedure in APBS (Baker et al. 2001). Accordingly, the MscL protein-lipid system has three dielectric regions: the high dielectric solvent exterior (80.0), the intermediate dielectric protein interior (10.0), and the low dielectric interior of the membrane (2.0). The membrane environment is applied except for the hole in which the MscL is located. Test calculations indicate that for MscL, the results are not sensitive to the value of the protein dielectric used.

As tests of the electrostatic solvation protocol, we compare the computed polar solvation free energy of two systems to values obtained based on the standard solvation protocol in APBS using atomistic structures. One system is a short peptide (with 27 residues and a length of ~4 nm) from APBS, for which the computed polar transferring free energy from bulk water into the membrane is 127.1 and 120.9 KJ/mol for the atomistic protocol and current CG-based protocol, respectively. The second system is the closed state of *E. coli*-MscL, for which the computed polar transferring free energy with the atomistic and CG models are 61.8 and 70.9 KJ/mol, respectively. These two examples suggest that the Martini/NLPB combination appears to provide a reasonable estimate for the polar solvation effects.

The apolar solvation force calculations follow the very generic framework described in Wagoner and Baker (2006) where the solvent-accessible surface area contribution is supplemented with volume and dispersion integral terms. A 0.3 spline window is used for the spline-based molecular surface, and the coefficients for surface tension and the volume term of the apolar calculations are 0.0042 KJ/mol/Å² and 0.23 KJ/mol/Å³, respectively. It is noted that the lipid membrane is not present in apolar solvation force calculations, while in fact most of the residues of TM2 and some residues on TM1 are buried in the membrane. Therefore, the apolar solvation forces on the outer particles of the transmembrane helices are not included in the continuum mechanics calculation (the interactions between these particles and lipids are included explicitly as discussed above). Care is taken to identify these particles based on the coordinates and van der Waals radius of each CG particle in the transmembrane helices.

Besides the outer particles of the transmembrane helices buried in lipids, the inner pore constriction at the closed state for the wild-type (WT) MscL is actually not hydrated at the first stage of gating until the pore gets large enough (> 0.45 nm (Beckstein et al. 2001)). In the above calculations, the channel pore is assumed to be already wetted at the beginning, which is probably more close to the gain-of-function (GOF) mutant MscL (e.g., the V23 T GOF mutant (Anishkin et al. 2005)). The hydration of the channel pore depends on the effective pore radius, and since the pore has different radii along the symmetry axis, at the beginning, only part of the pore surface is hydrated and the rest is not. In the closed state of MscL, about half of the TM1 bundle's inner surface is exposed to solvent and the lower half is not. When considering the hydration process depending on the effective pore radius as a more realistic model for the WT MscL, similar to those outer residues buried in lipids, the apolar solvation forces on the inner hydrophobic constriction region are not included in the continuum mechanics calculations. This is likely a reasonable approximation in the context of continuum modeling, although the hydration of MscL is a complex case that requires thorough MD studies (Beckstein et al. 2001; Beckstein and Sansom 2004).

In the following context, when considering hydrophobic interactions between water and the channel, the effect of gradual hydration is not included first; and the pore radius depended hydration process is considered in the section of “[Different Pathways for the GOF \(Gain-of-Function\) mutant and WT \(Wild-Type\) MscL.](#)”

During the continuum calculations, the positions of the CG model are constantly updated based on the coordinates of the “chemical nodes” of continuum mechanics MscL model, while the particle type, radius, charge, and interaction parameters are retained. And solvation forces of the constantly updated CG model calculated through the above protocols in turn participate in the FEM calculations. Solvation effects on the lipid membrane are neglected.

A Force-Based CM/CS Co-simulation Protocol

Shown in Fig. 13 is the assembled continuum model of *E. coli*-MscL under equibiaxial tension with zoomed-in views near the protein (the 10-petal lipid hole surface and the *E. coli*-MscL). Four-node tetrahedron finite elements are used

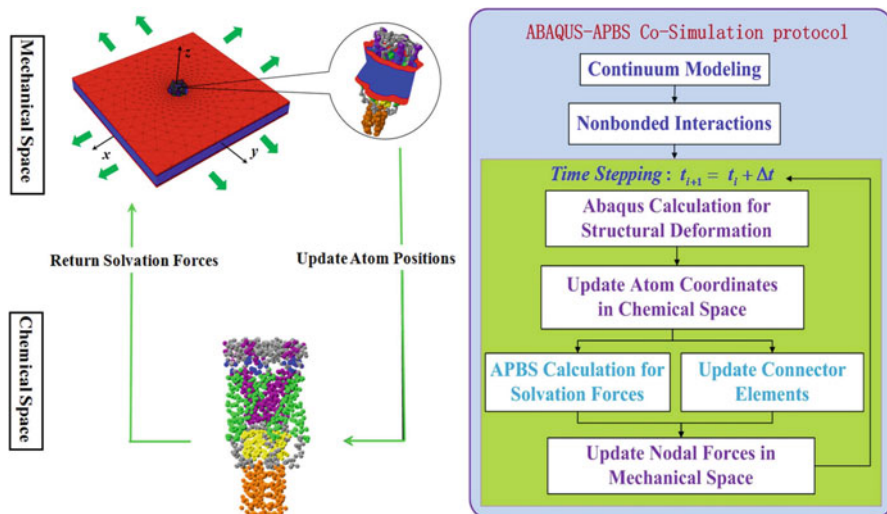


Fig. 13 Force-based ABAQUS-APBS co-simulation protocol: schematic (with a zoomed-in view near the protein) and flowchart representation. Zhu et al. (2016), reprinted with permission of Springer

to mesh both the helices and the lipid membrane with the commercial package ABAQUS (2011). Each chain of the continuum MscL model consists of 2,685 nodes and 9,699 finite elements. The lipid bilayer incorporates 2,412 nodes and 10,680 elements, with the mesh gradually more refined toward the boundary of the inner hole where it interacts with the protein extensively. As a reference calculation, MscL gating is modeled without including the solvation effects. Through ABAQUS simulation, the structural deformation of the lipid-protein system is explicitly calculated in response to an external load, where the lipid membrane is stretched (Fig. 13) by applying equi-biaxial displacement on its outer boundary (relevant to osmotic pressure).

To explicitly explore the solvation effect on MscL gating pattern, continuum solvation forces (electrostatic or apolar) need to be integrated into the continuum model that solves for lipid-protein system deformation under external force. This requires an iterative procedure that alternates between CM simulation and CS calculations (illustrated in Fig. 13):

1. Set up nonbonded interactions within MscL and those between MscL and lipid membrane.
2. A small equi-biaxial displacement on the lipid's outer boundary is applied, and a deformed conformation of the system is obtained by finite element analysis using ABAQUS.
3. The positions of all "chemical nodes" of MscL molecule are extracted from the deformed mesh, and then continuum solvation forces are calculated by APBS.

4. At the same time, the connector elements (interactions within MscL) are updated: new connector elements are added for pairs of two “chemical nodes” that are originally remote but come close during the gating process. The forces exerted by the connector element are directly applied to the corresponding nodes in the continuum model without interrupting the integrity of the element community. Since the gating mechanism of MscL is a gradually expanding process, there are only a few of newly applied connector elements (forces).
5. The solvation forces on the chemical nodes and forces from the newly emerged connector elements are included in the FEM simulation, and a new deformed structure is obtained in the next time step, leading to a new iteration.
6. Dozens of iterations (steps) are carried out until the channel is fully opened, i.e., when the channel pore radius of the continuum model reaches the experimentally estimated value (~ 19 Å) for the fully opened state (Sukharev et al. 2001b). Note that the solvation contribution includes both electrostatic and apolar components, and their effects are explored separately below. With different salt concentrations, the procedure above can be repeated to validate Sukharev et al.’s experimental observation (Sukharev et al. 1999) that there is no significant change in the gating pattern for salt concentrations between 0.05 and 1.0 M.

For a typical tension simulation of the gating of *E. coli*-MscL, the computational time is ~ 8 h on a Dell workstation with 3.2 GHz CPU and 4 Gb RAM, highlighting the efficiency of the MDeFEM framework compared to atomistic MD simulations, which take from days to months for large membrane protein systems.

Results and Discussion

In this section, we compare the gating pathways of three systems: (1) a pure CM model with isotropic membrane properties (i.e., in absence of membrane fluidity and solvation effects), (2) a CM model with membrane fluidity but lacking solvation effects, and (3) a complete model that incorporates the electrostatic/apolar solvation effect through the ABAQUS-APBS co-simulation protocol developed above. Simulation results are compared to structural models, previous all-atom simulations, as well as available experimental results.

Gating Pathway of MscL Without Membrane Fluidity and Solvation Effects

As a reference model, same as our previous MDeFEM study (Chen et al. 2008; Tang et al. 2008), we first treat each layer of the lipid as an isotropic slab without any explicit solvation effects; this is referred to as the preliminary model below. Shown in Fig. 14 are several snapshots of *E. coli*-MscL during the tension-driven gating process in comparison to the structural model of Sukharev and Anishkin (2004) at closed, half-, and fully opened states. Not surprisingly, the transmembrane region (TM1 and TM2 helices) exhibits the most striking conformational changes, in which both TM1 and TM2 helices shift radially away from

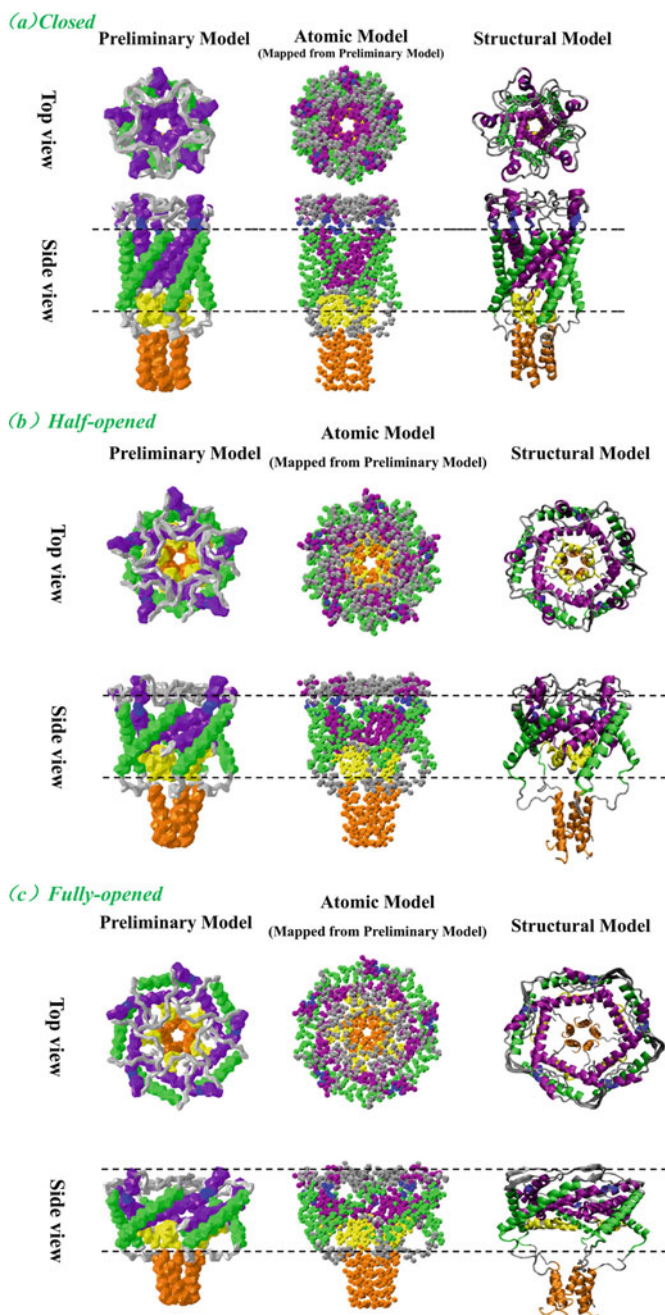


Fig. 14 Gating pathway of *E. coli*-MscL at closed (a), half- (b), and fully opened (c) states. Configurations are illustrated by continuum model in mechanical space and atomic model (VDW) in chemical space with comparison to the structural models (Sukharev and Anishkin 2004). The dashed lines indicate the approximate location of membrane/water interface. Zhu et al. (2016), reprinted with permission of Springer

the fivefold symmetry axis. Gating is primarily realized through the nonbonded interactions between transmembrane helices and the lipid membrane, in which the pore enclosed by TM1s is pulled open. The S1 bundle expands in the radial direction following the path of TM1 though it does not open as large as the TM1 bundle. At the current maximum lipid strain ($\epsilon_m=21\%$), no disassembly of S3 helices is observed; although the top region of the S3 bundle is expanded slightly, the lower ends remain assembled. This observation is in agreement with the revised version of the structural model (Sukharev and Anishkin 2004), as opposed to the earlier one (Sukharev et al. 2001b). Beside radial expansion of the TM helices, visible shrinking in the height of MscL is also observed, which is correlated with the significant titling of the helices and uplifting of the S3 bundle. Similar to previous studies (Tang et al. 2006, 2008), the loops and transmembrane helices are considerably stretched and bent during the gating process to maintain mechanical equilibrium; these features may be verified from experimental studies with sufficient resolution. On a qualitative level, the current MDeFEM results are in good agreement with the structural models (Sukharev and Anishkin 2004), regarding both the orientation and displacement of the helices and loops (Fig. 14). The only exception regards the periplasmic loops that link TM1 and TM2 helices; they remain well packed in our simulations but expand radially in the structural model of Sukharev and Anishkin (2004) along with the tips of the transmembrane helices.

The conductance of MscL is directly correlated with the size of the gate, and it is commonly believed that the TM1 helices in the core of the transmembrane bundle constitute the most important gate of the channel (Sukharev and Anishkin 2004; Tang et al. 2006), which is pentagon-shaped (insert of Fig. 15) when projected onto the x - y (membrane) plane. To characterize the pore size, we define an effective pore radius (denoted as r) as the radius of a circle that has the same area as the pentagon-shaped TM1 pore (see insert of Fig. 15). To be consistent with previous studies (Chen et al. 2008; Tang et al. 2008), the area of the pentagon-shaped TM1 pore is calculated as that surrounded by the five TM1 helical axes. From the closed state to the fully opened state, the effective radius increases from ~ 6.5 Å to ~ 19 Å, echoing the structural model (Sukharev and Anishkin 2004). Considering the pore's irregular inner surface, the actual pore's radius ranges from <1 Å to ~ 14 Å which is consistent with the estimation from fluorescence resonance energy transfer (FRET) experiments (Corry et al. 2010; Wang et al. 2014). For the preliminary model, the percentage of increment of the effective pore radius and the actual pore radius is depicted in Fig. 15 as a function of membrane strain (the inverted triangle curve). Due to the more detailed representation of MscL and the nonbonded interactions, the relationship between the membrane strain and the lipid cavity expansion is much less linear as found in the previous study (the triangle curve (Tang et al. 2008)). As mentioned in the Introduction, the surface-based nonbonded interaction model used in the previous study is oversimplified, and most importantly, it neglects all the nonbonded interactions involving the loops that connect helices. The current work demonstrates that as the membrane strain increases, the pore size of MscL is firstly stabilized at a small value and then experiences a rapid increase leading to pore opening when the membrane strain becomes sufficiently large.

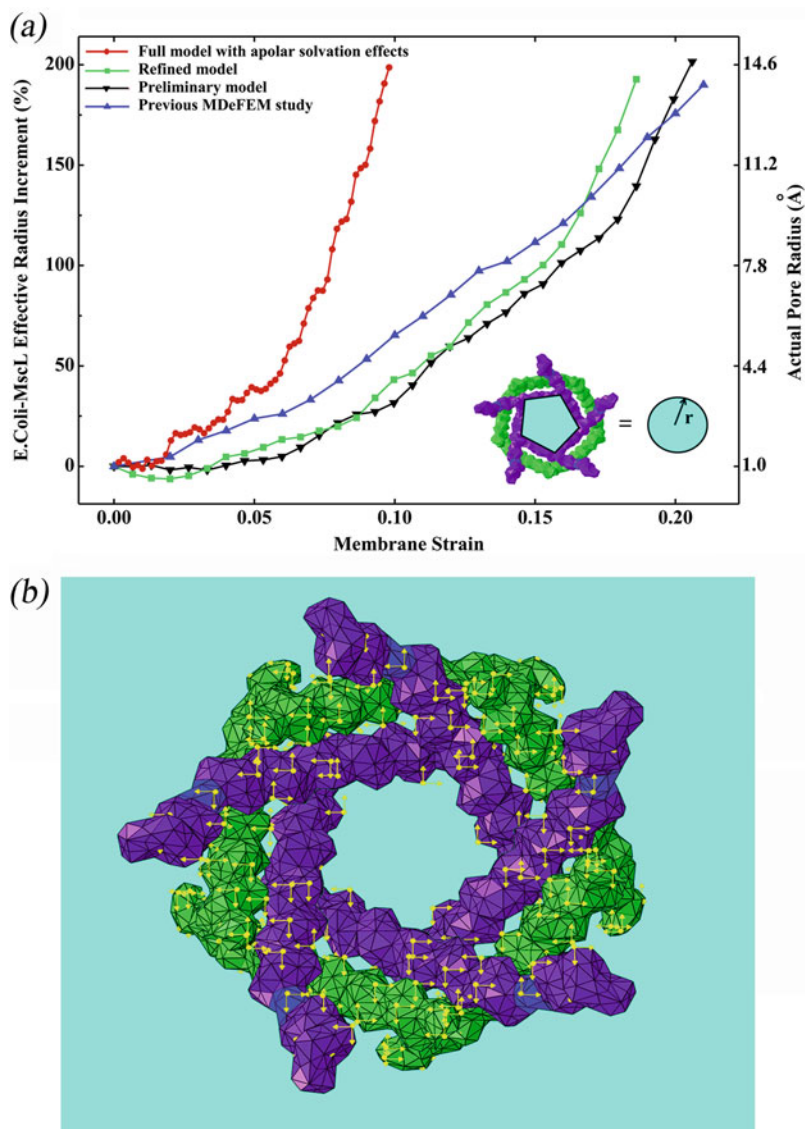


Fig. 15 Predicted behavior of MscL upon equi-biaxial tension (a) and a representative illustration of the apolar solvation forces (directions only) during MscL gating (b). Evolution of the effective pore radius and approximate actual radius of MscL is shown as a function of membrane strain. The results obtained currently are compared with those of the previous MDeFEM model (Tang et al. 2008). The refined model includes the effect of the lipid's fluidity, and the full model further includes the apolar solvation effects. Zhu et al. (2016), reprinted with permission of Springer

This observation is in qualitative agreement with the patch-clamp experiment of Sukharev et al. (1999), in which the tension-activation data was fitted to a Boltzmann form.

Gating Pathway of MscL with Fluidic Membrane and Apolar Solvation Effects

As emphasized above, two major limitations of our previous work (Chen et al. 2008; Tang et al. 2008) and the preliminary model concern the lack of membrane fluidity and solvation effects. In the following, we first include the fluidic properties of the membrane, leading to the “refined model.” Next, we further include solvation effects, leading to the “full model.” Since we expect that the interactions between the apolar surface of the MscL pore and the solvent make a dominant contribution to the gating process, in this section we only include effects of apolar solvation; the relative importance of electrostatic and apolar solvation forces will be discussed in the next section.

With the fluidic lipid model and continuum apolar solvation model employed successively, new effective pore radius evolution curves are obtained and shown in Fig. 15 (square curve for the refined model and circle curve for the full model with apolar solvation). The overall trend of the gating pathway remains the same as that of the preliminary model (inverted triangle curve) in that the pore size is firstly stabilized at a small value and then experiences a rapid increase as the lipid strain further increases. Compared to the preliminary model, the most notable difference lies in the lipid membrane strain required for fully opening MscL. With the more realistic (orthotropic) model of the lipid, the membrane strain for full opening of MscL is reduced slightly ($\sim 10\%$); with the apolar solvation included, the gating strain is further reduced by as much as $\sim 50\%$. Figure 15b shows the solvation forces (directions only) on the “hydrophobic chemical nodes (residues)” on TM1s and TM2s in one step of the simulation of MscL gating. Since most of the residues on the transmembrane helices are hydrophobic rather than hydrophilic (Fig. 10b), the interactions between water- and solvent-exposed residues are expected to push the channel outward, providing another driving force (in addition to the membrane tension) for the gating transition of MscL.

Take the hydrophobic residues on the constriction area as an example. In the closed state, these residues are very close, and hydrophobic confinement is thermodynamically favorable since exposure of these hydrophobic residues is energetically unfavorable. But when the pore constriction is opened wide enough (mainly by membrane stretch), water molecules will be “driven” to fill in the constriction space despite the fact that exposure of hydrophobic residues on pore constriction area is still energetically unfavorable. This “driving” factor that compel the water to fill into the opened space is probably the system’s internal pressure, change of which is usually not considered in MD simulations. Thus the “unfavorable” energy required to expose the hydrophobic residues on the pore constriction is compensated by drop of the internal pressure which may be very small on the scope of the whole system but critical for MscL gating. After that, the repulsion interaction between water and these hydrophobic residues is helping channel gating.

It is known that the gating of MscL can be triggered solely by membrane stretching, although the importance of apolar solvation to gating has only been indirectly probed. The release of content through *Tb*-MscL has been studied by a pioneering coarse-grained MD simulation work (Louhivuori et al. 2010) which,

however, did not explicitly elucidate the importance of the hydrophobic effect in MscL gating. A molecular dynamics study on *Tb*-MscL (Jeon and Voth 2008) observed that the water chain formation across the channel pore took place at the same time as the channel pore radius increased to a certain value, and they suggested that the two processes may provide positive feedback to each other. The results we present here provide direct evidence for this point. Though the importance of hydrophobic interactions in channel gating has been demonstrated previously by hydrophilic mutations or molecular dynamics studies (Anishkin et al. 2005; Jeon and Voth 2008), this work, to the best of our knowledge, represents the first explicit evaluation of the hydrophobic contribution.

Another work that concerned hydration process of MscL gating is conducted by Anishkin et al. (2010) who focused on the hydrophilic interaction between water and buried hydrophilic residues instead of hydrophobic interaction as we present in this work. Based on the analysis of the hydration energy of the pore constriction and the fact that hydrophilic mutations in the pore constriction area make the channel opening more easy, a conclusion is drawn that the process of the glycine (hydrophilic residue) exposure with pore expansion and their favorable hydration create disjoining pressure that assists opening (Anishkin et al. 2010).

Based on Anishkin et al.'s hydration energy analysis and the analysis of the results in our present work, detailed effects of the hydrophilic or hydrophobic residues on channel gating are elucidated as follows. In the process of the exposure of initially buried hydrophilic residues in the pore constriction area or buried in other area, due to their favorable hydration, disjoining pressure will be generated that assists conducting as suggested by Anishkin et al. (2010). But once these hydrophilic residues are exposed, their interaction with water molecules will resist opening of the channel based on our analysis primarily due to the hydrogen bonding between the polar residue and water molecules. On the other hand, the effects of the hydrophobic residues also act in a similar but contrary way. Before exposure of the hydrophobic residues in the hydrophobic constriction area, these hydrophobic residues are closely packed and the interaction within helps the channel keep closed. But once these hydrophobic residues are exposed, their interaction with water molecules will facilitate opening of the channel primarily due to the repulsion interaction between water and nonpolar residues. Collectively, that is, the effects of the hydrophobic or hydrophilic residues are reversed before and after exposure to water. These mechanisms may be further verified by future experimental and simulation studies.

Shown in Fig. 16 are the configurations at half-/fully opened states of the refined and full models in comparison to the structural model of Sukharev and Anishkin (2004). As with the preliminary model, the conformational transitions observed with the refined or full model are also in qualitative agreement with the structural model. Closer inspection shows that opening of the S1 bundle with the full model is much more significant and the pore size surrounded (not enclosed) by S1 helices are also comparable with that of the TM1 pore (Fig. 16). Note that the position of the S1 domain in the opened state may be different from the real case for biological MscL molecule since the S1 domain of the initial MscL structure (Sukharev and Anishkin

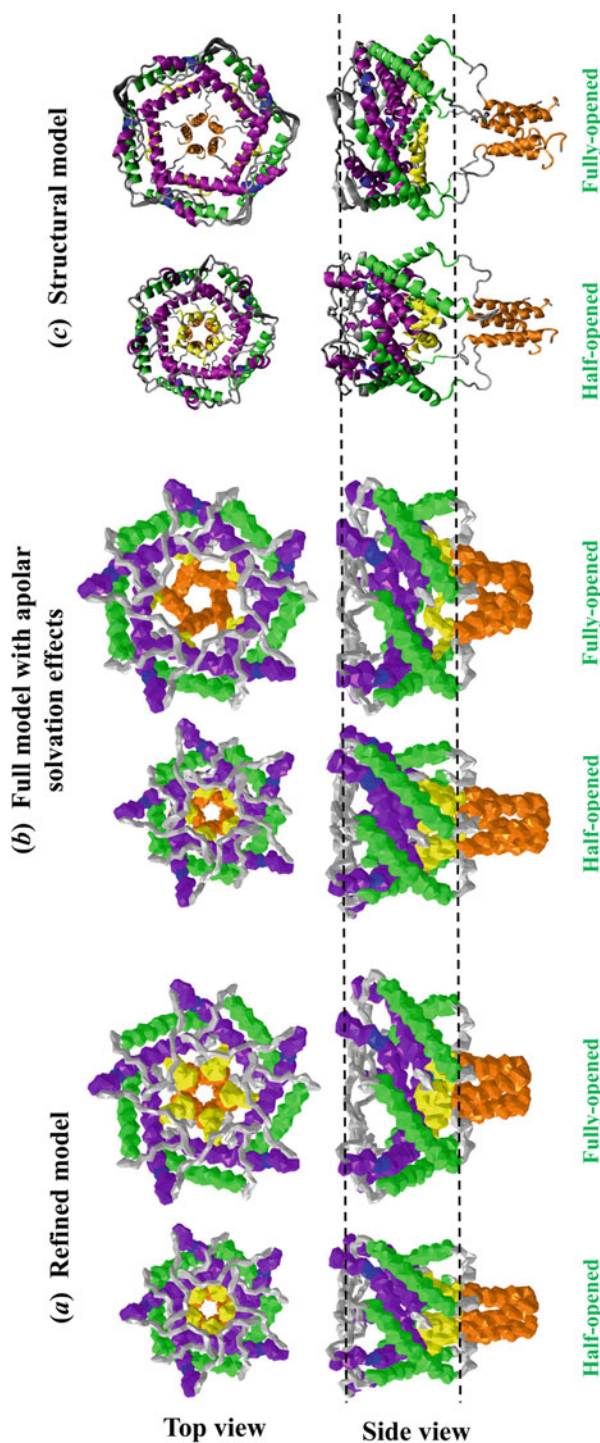


Fig. 16 MscL gating pathway of the refined model (a) and the full model (b). Results at half- and fully opened states are shown with comparison to the structural model (c) (Sukharev and Amishkin 2004). The dashed lines indicate the approximate location of membrane/water interface. Here, the refined model takes into account the fluidity of the lipid membrane, and the full model further includes the apolar solvation effects. Zhu et al. (2016), reprinted with permission of Springer

2004) we use may not be appropriately positioned (Iscla et al. 2008). The S3 bundle also separates more in the full model especially for the upper ends though its lower ends remain assembled together. Similar to the preliminary model, the loops linking TM1-TM2 remain essentially unseparated, and this result is consistent with that of the SMD study (Gullingsrud and Schulten 2003) in which the periplasmic loop region in all simulations remained stable. Experimentally, it was found that the channel remains mechanosensitive even with the external loops cleaved (Ajouz et al. 2000); thus, the precise functions of these loops remain unclear. The stable behavior of the periplasmic loops during the simulated gating process suggests that they may function as a filter screen to prevent large molecules from entering and occluding the channel while at the same time prolonging channel opening.

Respective Effects of Electrostatic and Apolar Solvation Forces

Previous patch-clamp experiments (Sukharev et al. 1999) showed that there was no significant change in the MscL gating pattern in the range of salt concentration between 0.05 and 1 M. This observation suggests that electrostatic interactions are unlikely to dominate in MscL gating. To further explore this hypothesis quantitatively, we simulate MscL gating at several ion concentrations (0.05 M and 2.0 M) in the framework of continuum solvation.

As shown in Fig. 17, there are minor differences between the effective pore radius evolution curves of the refined model, the full model with electrostatic solvation contributions with ion concentration of 0.05 M or 2.0 M, respectively. Configurations from the full model with electrostatic solvation effects (not shown) are also very close to those from the refined model (ion concentration, 0.00 M) shown in Fig. 16. Compared to the refined model where no solvation contributions are included, the effective pore radius evolution does not change much when only the electrostatic solvation forces are considered. The apolar solvation forces, on the other hand, contribute significantly to the gating process (circle curve in Fig. 17 or Fig. 15) as described in the above section.

Different Pathways for the GOF (Gain-of-Function) mutant and WT (Wild-Type) MscL

The hydrophobic core of the TM1 bundle of WT MscL appears to be dehydrated according to the EPR data and molecular dynamics results (Gullingsrud and Schulten 2003; Perozo et al. 2002a). The energy landscape for WT MscL indicates that the major energy cost for MscL opening is between the closed state and the substate $S_{0.13}$ (Anishkin et al. 2005) (0.13 means the relative conductance to the fully opened state) and this substate is believed to be already well hydrated (Anishkin et al. 2005). Experiments showed that hydrophilic substitutions in the hydrophobic restriction of TM1s led to a reduction of the gating tension, while a more hydrophobic substitution resulted in a channel that requires a greater tension to open (Anishkin et al. 2005). Based on these data, one could associate the passage of the main energy barrier with the hydration of the largely hydrophobic pore (Anishkin et al.

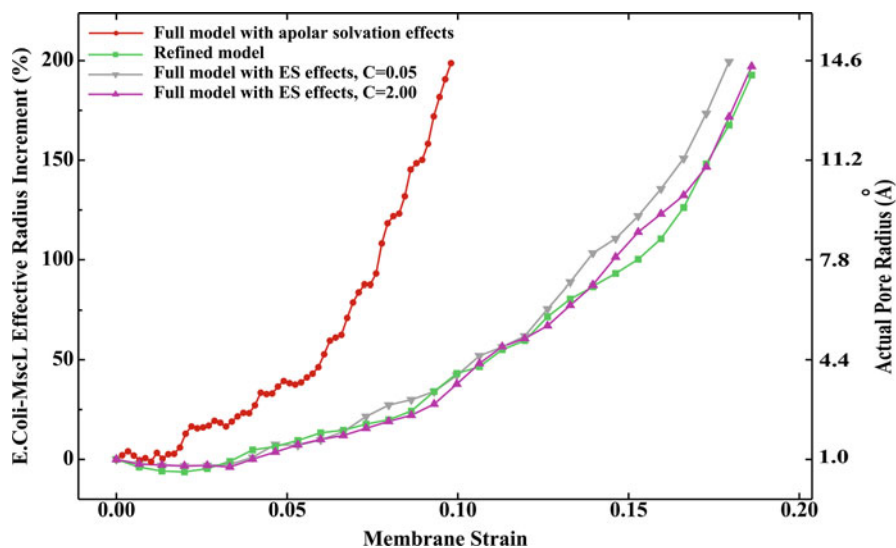


Fig. 17 Evolution of the effective pore radius of MscL versus membrane strain for different models. The different effects of the apolar and electrostatic solvation forces on the gating of *E. coli* MscL are obvious. The full model includes the effects of the electrostatic solvation (ES) (with ion concentration C) or apolar solvation (AS) forces. Zhu et al. (2016), reprinted with permission of Springer

2005), which is a prerequisite for conduction. The hydrophilic substitution in the hydrophobic restriction of TM1s may disrupt this initial hydrophobic restriction, leading to facile hydration of the pore and reduction of the first free energy barrier.

In the present simulation, the hydration of the channel pore depends on the effective pore radius as described in the “Models and methods” section. The threshold radius for hydration is taken to be 0.45 nm (Beckstein et al. 2001) as indicated in the right bottom inset of Fig. 18. In the closed state of MscL, about half of the TM1’s inner surface is exposed to solvent and the lower half is not. When considering the hydration process here, accordingly, the apolar solvation forces on the inner hydrophobic constriction region are not included in the continuum mechanics calculations.

Figure 18 shows the effective radius evolution curves from full model simulations (including apolar solvation effects) for the WT and the GOF mutant MscL (hypothetical hydrophilic substitution in the hydrophobic constriction of TM1s). A rightward shift of the effective radius evolution curve is observed for the WT MscL as compared to the GOF mutant. This result is in agreement with experimental observation that an initially better hydrated state (GOF mutant MscL caused by hydrophilic substitution) is easier to open (Anishkin et al. 2005). We note that the TM1 helices are not fully hydrated during the simulation until the membrane strain gets to as large as $\sim 2/3$ of the strain required for full gating (indicated by the dashed

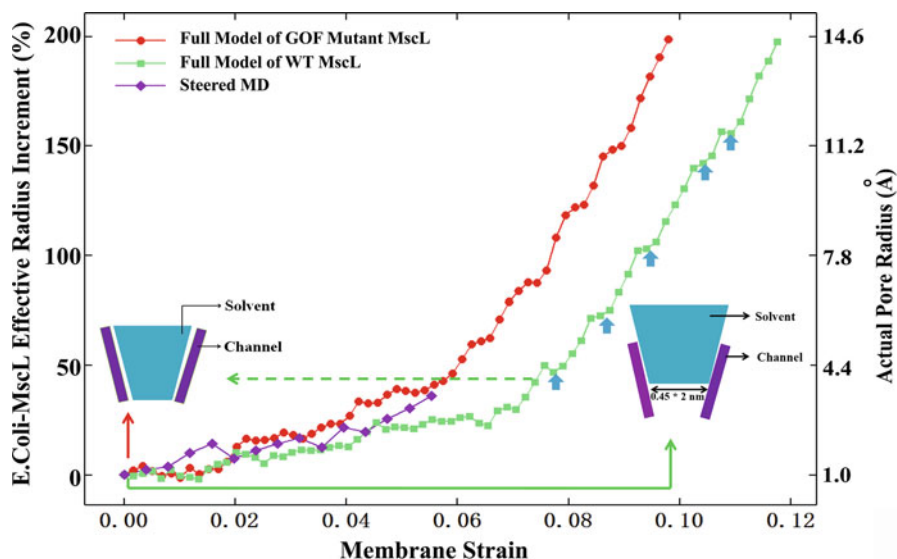


Fig. 18 Effective radius evolution of simulations for the GOF mutant and WT MscL. The GOF mutant MscL has a hydrophilic substitution in its hydrophobic constriction causing full hydration in the initial state, while the initial state of the WT MscL is only half hydrated. The WT channel won't get fully hydrated until the membrane strain reaches $\sim 2/3$ of the gating strain (indicated by the dashed line) which agrees well with Sukharev et al.'s experimental study (Sukharev et al. 1999). Zhu et al. (2016), reprinted with permission of Springer

line in Fig. 18), suggesting that MscL is nonconductive until two thirds of its gating tension is reached. This result is in agreement with Sukharev et al.'s experimental observation (Sukharev et al. 1999) that the open probability remains zero until the membrane tension reaches ~ 10 dyne/m and the gating tension is about 15 dyne/m. Thus, when the membrane strain reaches half of the gating strain for MscL, the mechanosensitive channel of small conductance (MscS) reaches its fully conducting and opened state (Martinac and Kloda 2003), while MscL is not yet conductive, emphasizing MscL's role of a "safety valve" in prokaryotes as a last-ditch effort for survival (Berrier et al. 1996).

The effective radius evolution of these two full models is compared to that in the steered molecular dynamics (SMD) study (Gullingsrud and Schulten 2003) (rhombus curve in Fig. 18). The maximum membrane strain in SMD is estimated to be $\sim 5.4\%$, and we assume that its increment is proportional to the time step used in SMD. Although the channel is far from fully open in the nanosecond SMD simulation, as shown in Fig. 18, the effective pore radius evolution in SMD is in good agreement with that of the continuum full model at small strain.

The final lipid strain (ϵ_m) of the full model for the complete opening of MscL is $\sim 11.76\%$ (gating strain). The strain in the direction normal to the membrane is (Tang et al. 2006)

$$\varepsilon_z = \frac{2\bar{v}_l}{1 - \bar{v}_l} \varepsilon_m \quad (5)$$

which leads to a reduction in the membrane thickness of

$$\Delta h = \varepsilon_z h \quad (6)$$

where $\bar{v}_l = 0.5$ is the equivalent Poisson's ratio of the orthotropic lipid membrane model and $h = 35\text{\AA}$ is the initial thickness of the membrane. With the membrane strain of 11.76%, the thickness is reduced from 35\AA to 28.77\AA , which is a 23.5% change. This value is in close agreement with the 20% reduction in the thickness of membrane spanning part of MscL measured by experiment (Perozo et al. 2002a) since the flatness of MscL is directly caused by thinning of the lipid membrane during gating.

Another prominent feature of the full model's gating pathway is the stepwise nature of the pore radius evolution curve, as shown in Fig. 18 the circle curve or the square curve. The initial survey of MscL in liposome patches (Sukharev et al. 1999) has recorded three short-lived subconducting states where the pore size stabilizes even in the presence of tension, and a later study of Chiang et al. (2004) identified nine subconducting states. In the present study, when the membrane strain is below $\sim 2/3$ of the gating strain, the channel is not fully hydrated and remains nonconducting. Afterward, the channel becomes conductive, and the effective pore radius increases very rapidly with the membrane strain. A closer inspection of the data shows some plateaus (marked with arrows in Fig. 18) of the curve where the pore size stabilizes. Although the current continuum model may not be sufficient enough to capture some delicate molecular phenomena, such as the subconducting states of MscL gating, nevertheless, the results here may imply that the subconducting states may be in part due to the intricate "overall" mechanical interactions of the multiple components of the system, which complement their biophysical functions.

The actual pore radius is about 4.75\AA for the first substate in Fig. 18, and the radius of the inner pore surface in the open state is $\sim 14\text{\AA}$, which is consistent with the estimation from fluorescence resonance energy transfer (FRET) experiments (Corry et al. 2010; Wang et al. 2014). Assuming that the conductance of the MscL pore is proportional to its size (area enclosed by the inner surfaces of TM1s) (Steinbacher et al. 2007), the relative conductance in the present study can be estimated as the ratio of the pore area to that of the full open state. For example, the relative conductance for the first plateaus in present work would be $(4.75/14)^2 = 0.115$, which is in good agreement of the experimental estimate of 0.13 for the first subconducting state (Anishkin et al. 2005). Their proposal that the main energy costing substate is well hydrated is also consistent with results shown in Fig. 18.

Analytical Effort

An important goal of this section is to establish a closed-form and simple analytical model (Chen et al. 2008), as an alternative approach that can capture the most

essential features of MscL gating that might be broadly applicable to MS channels, such as the different contributions of solvation and membrane stretch to gating. An analytical model that couples MscL to both lipid membrane stretch (outer boundary) and apolar solvation effects (inner boundary) is developed as follows.

For the square lipid membrane (Fig. 19a) with a length of $2l$ (400 Å) and a central cavity radius of c , its outer boundary pressure is $-p_{\text{lipid}}$, and the interface pressure between the lipid and MscL is $-p_{\text{lipid-MscL}}$. The whole lipid membrane is treated as an isotropic plate. Since the in-plane fluidity of the lipid membrane contributes only slightly (less than 10%) to the gating process, this treatment is considered a reasonable approximation. During gating, the deformation of the membrane cavity is mainly transferred to the closest TM2 helices via nonbonded interactions in the radial direction. The nonlinear interaction pressure-distance relationship is analogous to a nonlinear elastic medium between the lipid cavity and the TM2 bundle. After the TM2 helices are pulled open, the TM1/TM2 nonbonded interactions (another effective nonlinear medium) may perturb the MS channel radius, which is enclosed by the five TM1 helices. Therefore, a simple analytical model can be established in which the details of protein structures are ignored and the nonbonded interactions are described by effective elastic media. A schematic of such plane stress effective continuum medium model (ECMM) is given in Fig. 19b with *E. coli*-MscL as an example. The inner effective annular medium I accounts for the TM1-TM1 interactions in hoop direction and TM1-TM2 interactions in radial direction, and the outer continuum medium II incorporates TM2-TM2 interactions in hoop direction and TM2-lipid interactions in radial direction. The inner radius, interface radius, and outer radius of the ECMM are denoted by a , b , and c , respectively. Here, a is the effective radius of the closed MscL (consistent with the definition in previous sections), which corresponds to the smallest “through” capacity of the TM1 bundle; b is defined similarly for the TM2 bundle; and c is the interface radius between MscL/lipid cavity. From the closed homology structure of *E. coli*-MscL, a , b , and c are equal to 6.5 Å, 17 Å, and 22 Å, respectively.

The outer boundary of MscL ECMM (Fig. 19b) is coupled to the inner cavity of the lipid model (Fig. 19a), i.e., $\sigma_r^{\text{II}}(c) = -p_{\text{lipid-MscL}}$, $u_r^{\text{II}}(c) = u_r(c)$. Continuity of radial stress and displacement at the interface between medium I and II ($r = b$) requires $\sigma_r^{\text{I}}(b) = \sigma_r^{\text{II}}(b)$, $u_r^{\text{I}}(b) = u_r^{\text{II}}(b)$. The inner boundary pressure of medium I is, $\sigma_r^{\text{I}}(a) = -p_{\text{water}}$, from apolar solvation effects. Following the theoretical analysis in Chen et al. (2008) but considering the coupling between MscL ECMM and the lipid model and the coupling between MscL ECMM and the solvation contribution, a closed-form solution of the MscL pore radius increment can be obtained as

$$\Delta a = 0.085 p_{\text{water}} + 60.205 \varepsilon_m \quad (7)$$

where the units for p_{water} and Δa are MPa and Å, respectively. On the right of Eq. (7), the first term is the contribution of apolar solvation effects to MscL gating and the second the contribution of membrane stretch. The two constants before p_{water} and ε_m depend on the elastic properties of medium I and II (Chen et al. 2008). The value of p_{water} depends on the exposed residues of the TM1 bundle and the state of

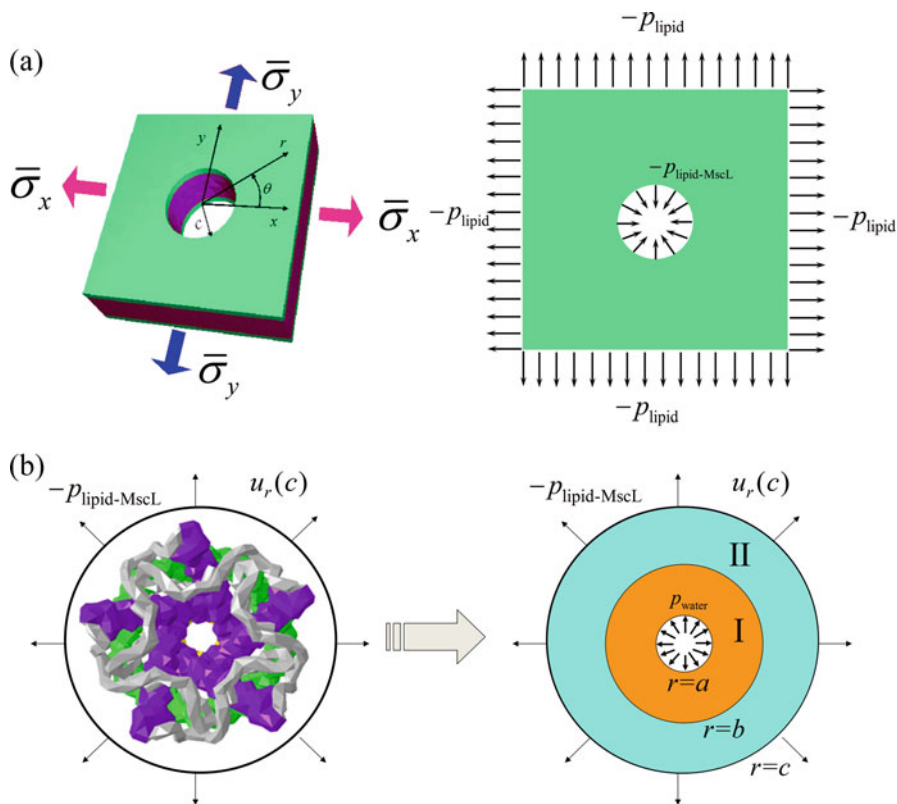


Fig. 19 Schematic of the linear lipid membrane model (a) and the effective continuum medium model (ECMM) of MscL (b)

the hydrophobic constriction. At the initial state, $p_{water} = 0$. Averagely, based on the continuum solvation forces calculated in the above sections and the exposed pore surface area, $p_{water} \approx 68$ MPa. Accordingly, the final membrane strain to fully open the channel is 11.16%, which is in close agreement with that (11.76%) calculated by the coupled CM/CS MDeFEM approach.

Limitations of the Current Implementation and Future Directions

Despite our tremendous endeavor and that the current model has been able to help us gain some useful perspectives for channel gating, there are considerable limitations for the present coupled continuum mechanical-continuum solvation approach, and the room for future improvement is still large.

One of the major limitations concerns the atomic structure of *E. coli*-MscL whose S1 domain may need revision (Iscla et al. 2008) based on the revised crystal

structure of *Tb*-MscL (Steinbacher et al. 2007). And the crystal structure of *Tb* MscL could also be subjected to the above developed continuum-based approach to study the effects of the S1 domain on the gating behavior. Through these improvements, the proposed facilitation effect of the S1 domain on gating due to its close interaction to the membrane (Iscla and Blount 2012; Vanegas and Arroyo 2014) could be verified and more specifically quantified.

Another limitation of the current work is the treatment of the lipid membrane and its surface-based interaction to the channel. The particle-to-particle lipid-channel interaction force is averaged to the surfaces of them; thus, the heterogeneous binding of the lipid to the channel is lost. Interfacial lipid-channel interactions are exceedingly complicated due to hydrophobic mismatch, electrostatic effects, dynamic nature of the lipid environment, hydrogen bonding, variation during conformational transition, asymmetry, localization, etc. (Argudo et al. 2016; Powl et al. 2003, 2005a, 2008; Vanegas and Arroyo 2014). While emerging evidence has shown the importance of the interaction of the membrane with membrane channels for many biological processes, it is still difficult to elucidate these interactions in a detailed and comprehensive manner both experimentally and computationally. These characteristics of lipid-channel interactions impose great challenge for continuum modeling of the lipid membrane and its interaction to the channel. For future work, one may consider developing particle-to-particle-based interactions and geometry refinement of the lipid membrane (especially for the lipid closely around the channel), which could be included in a similar way as how we treat the channel in the current work. And solvation forces can then be applied to the lipid membrane model as well to mimic a hydration environment for the lipid model. Thereafter, this well-established lipid model can be exploited to explore the contribution or mechanism of different lipid-channel interactions for channel gating in future work. For example, how and how much does the “anchor” effect between S1 domain and the lipid membrane affects gating (Iscla and Blount 2012)? How and how much does the electrostatic interaction between the lipid head group and the charged transmembrane helix residues affect channel gating? Is the distortion of the lipid membrane around the channel helping channel gating, or is it just a spontaneous adaptation behavior for the system to reach a low energy state? Will the distortion of the membrane around the channel disappear or diminish during gating? Besides, this new lipid membrane is also much more reasonable for a study of the effect of negative or positive hydrophobic mismatch (Perozo et al. 2002b), although the current lipid model can also catch some of the basic principle that, as an example, for the wild-type MscL full model with apolar solvation effects, increasing the hydrophobic (tail group) layer of the current lipid membrane model by 10% will lead to a 2.8% increase of gating tension for channel opening.

Moreover, some other subtle but potentially important molecular details such as inclusion of side chains in the FEM model or initial turgor pressure in living *E. coli* cells (Deng et al. 2011) may also play a part in the gating process and need further refinement in future continuum modeling. We hopefully expect that these

refinements will elaborately improve the accuracy of the continuum-based model and greatly expand the application scope and universality of this coupled continuum mechanics-continuum solvation approach.

Concluding Remarks

The analysis of MS channel gating provides us with an effective window to explore how mechanical stimuli induce adaptive cellular behaviors through protein structural transitions across different time and length scales. As shown by the results of the present work, their gating patterns highlight the roles of large-scale helical movements, pore hydration, and protein-lipid interactions during ion channel gating transitions. These processes occur on time and length scales that are too large to be studied directly by regular atomistic simulations. In this chapter, we have modified a molecular dynamics-decorated finite element (MDeFEM) method to incorporate key interaction pairs (e.g., Asp62-Arg131 and Phe68-Leu111) into a continuum mechanics model; this allows us to explore the gating pathway of MscS and how specific interactions and structural motifs impact the gating transition. Besides, a novel simulation protocol is developed that effectively integrates continuum solvation contributions (CS) into continuum mechanics (CM) calculations to study the gating pathway of MscL.

A complete gating transition trajectory of MscS from the closed to the open state along with partially open intermediates is obtained, and the open structure is close to the available structural model from crystallographic studies. It is observed that removing either the Asp62-Arg131 salt bridge or the Phe68-Leu111 nonpolar interaction leads to essentially nonconducting structures. The loop connecting TM2 (the second transmembrane helix) and TM3 is found to be essential for force transmission during gating, while the loop connecting TM1 and TM2 does not make any major contribution. Based on the different structural evolutions observed when the TM3 kink is treated as a loop or a helical segment, we propose that the helical propensity of the kink plays a central role in inactivation. Gating transition of MscS under different transmembrane voltages is also explored.

A novel computational framework is further developed by using MscL as a template. The continuum mechanics is closely coupled with the continuous solvent model. The influence of the solvent and the chemical coupling force is obtained by real-time iteration between the mechanical and chemical spaces. Compared to previous continuum mechanics studies, the present model is capable of capturing the most essential features of the gating process in a much more realistic fashion: due mainly to the apolar solvation contribution, the membrane tension for full opening of MscL is reduced substantially to the experimental measured range. A significant fraction ($\sim 2/3$) of the gating membrane strain is required to reach the first subconducting state of our model, which is featured with a relative conductance of 0.115 to the fully opened state. These trends agree well with experimental observations.

We expect that many of the mechanical principles discussed here are also likely to play a role in various membrane-mediated biomechanical processes. The successful application of the MDeFEM approach to MscS/MscL suggests similar studies of the growing families of sensory channels (Chen et al. 2015; Clapham 2003; Dhaka et al. 2006; Krishtal 2003; Pruitt et al. 2014) and their modulations by lipids, lipid-soluble factors, temperature, cell volume, and membrane tension. The approach is particularly powerful in cases that involve large length and long time scales, which are usually not easily accessible to standard particle-based simulations.

References

- ABAQUS 6.11 User's Manual. ABAQUS Inc., Providence, RI, (2011)
- B. Ajouz, C. Berrier, M. Besnard, B. Martinac, A. Ghazi, *J. Biol. Chem.* **275**, 1015 (2000)
- B. Akitake, A. Anishkin, S. Sukharev, *J. Gen. Physiol.* **125**, 143 (2005)
- B. Akitake, A. Anishkin, N. Liu, S. Sukharev, *Nat. Struct. Mol. Biol.* **14**, 1141 (2007)
- A. Anishkin, C. Kung, *Curr. Opin. Neurobiol.* **15**, 397 (2005)
- A. Anishkin, S. Sukharev, *Biophys. J.* **86**, 2883 (2004)
- A. Anishkin, S. Sukharev, *Channels (Austin)* **11**, 173 (2017)
- A. Anishkin, C.S. Chiang, S. Sukharev, *J. Gen. Physiol.* **125**, 155 (2005)
- A. Anishkin, B. Akitake, S. Sukharev, *Biophys. J.* **94**, 1252 (2008a)
- A. Anishkin, K. Kamaraju, S. Sukharev, *J. Gen. Physiol.* **132**, 67 (2008b)
- A. Anishkin, B. Akitake, K. Kamaraju, C. Chiang, S. Sukharev, *J. Phys. Condens. Matter* **22**, 454120 (2010)
- D. Argudo, N.P. Bethel, F.V. Marcoline, M. Grabe, *Biochim. Biophys. Acta Biomembr.* **1858**, 1619 (2016)
- N.A. Baker, D. Sept, S. Joseph, M.J. Holst, J.A. McCammon, *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10037 (2001)
- R.B. Bass, P. Strop, M. Barclay, D.C. Rees, *Science* **298**, 1582 (2002)
- N. Bavi, O. Bavi, M. Vossoughi, R. Naghdabadi, A.P. Hill, B. Martinac, Y. Jamali, *Channels (Austin)* **11**, 209, (2017)
- N. Bavi et al., *Nat. Commun.* **7**, 11984 (2016b)
- O. Beckstein, M.S. Sansom, *Phys. Biol.* **1**, 42 (2004)
- O. Beckstein, P.C. Biggin, M.S.P. Sansom, *J. Phys. Chem. B* **105**, 12902 (2001)
- V. Belyy, A. Anishkin, K. Kamaraju, N. Liu, S. Sukharev, *Nat. Struct. Mol. Biol.* **17**, 451 (2010)
- C. Berrier, M. Besnard, B. Ajouz, A. Coulombe, A. Ghazi, *J. Membr. Biol.* **151**, 175 (1996)
- R.B. Best, X. Zhu, J. Shim, P.E. Lopes, J. Mittal, M. Feig, A.D. Mackerell Jr., *J. Chem. Theory Comput.* **8**, 3257 (2012)
- M. Betanzos, C.S. Chiang, H.R. Guy, S. Sukharev, *Nat. Struct. Biol.* **9**, 704 (2002)
- H. Binder, K. Gawrisch, *J. Phys. Chem. B* **105**, 12378 (2001)
- P. Blount, M.J. Schroeder, C. Kung, *J. Biol. Chem.* **272**, 32150 (1997)
- I.R. Booth, M.D. Edwards, S. Black, U. Schumann, S. Miller, *Nat. Rev. Microbiol.* **5**, 431 (2007)
- K. Borngen, A.R. Battle, N. Moker, S. Morbach, K. Marin, B. Martinac, R. Kramer, *Biochim. Biophys. Acta* **1798**, 2141 (2010)
- B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus, *J. Comput. Chem.* **4**, 187 (1983)
- G.X. Cao, X. Chen, *J. Mater. Res.* **21**, 1048 (2006)
- G. Cao, X. Chen, *J. Mater. Res.* **21**, 1048 (2011)
- E. Chacon, P. Tarazona, F. Bresme, *J. Chem. Phys.* **143**, 034706 (2015)
- G. Chang, R.H. Spencer, A.T. Lee, M.T. Barclay, D.C. Rees, *Science* **282**, 2220 (1998)
- X. Chen, Y. Tang, G. Cao, *Proc. Inst. Mech. Eng. Part N* **219**, 73 (2006)

- X. Chen, Q. Cui, Y. Tang, J. Yoo, A. Yethiraj, *Biophys. J.* **95**, 563 (2008)
- B. Chen, B. Ji, H. Gao, *Annu. Rev. Biophys.* **44**, 1 (2015)
- C.S. Chiang, A. Anishkin, S. Sukharev, *Biophys. J.* **86**, 2846 (2004)
- D.E. Clapham, *Nature* **426**, 517 (2003)
- B. Corry, A.C. Hurst, P. Pal, T. Nomura, P. Rigby, B. Martinac, *J. Gen. Physiol.* **136**, 483 (2010)
- C. Cui, D.O. Smith, J. Adler, *J. Membr. Biol.* **144**, 31 (1995)
- M.E. Davis, J.A. Mccammon, *Chem. Rev.* **90**, 509 (1990)
- Y. Deng, M. Sun, J.W. Shaevitz, *Phys. Rev. Lett.* **107**, 158101 (2011)
- E. Deplazes, M. Louhivuori, D. Jayatilaka, S.J. Marrink, B. Corry, *PLoS Comput. Biol.* **8**, e1002683 (2012)
- L. Deseri, P. Pollaci, M. Zingales, K. Dayal, *J. Mech. Behav. Biomed. Mater.* **58**, 11 (2016)
- A. Dhaka, V. Viswanath, A. Patapoutian, *Annu. Rev. Neurosci.* **29**, 135 (2006)
- R.O. Dror, R.M. Dirks, J.P. Grossman, H. Xu, D.E. Shaw, *Annu. Rev. Biophys.* **41**, 429 (2012)
- M.D. Edwards, W. Bartlett, I.R. Booth, *Biophys. J.* **94**, 3003 (2008)
- R. Gamini, M. Sotomayor, C. Chipot, K. Schulten, *Biophys. J.* **101**, 80 (2011)
- M.A. Geeves, K.C. Holmes, *Annu. Rev. Biochem.* **68**, 687 (1999)
- M.A. Geeves, K.C. Holmes, *Adv. Protein Chem.* **71**, 161 (2005)
- J. Gullingsrud, K. Schulten, *Biophys. J.* **85**, 2087 (2003)
- J. Gullingsrud, K. Schulten, *Biophys. J.* **86**, 3496 (2004)
- O.P. Hamill, B. Martinac, *Physiol. Rev.* **81**, 685 (2001)
- E.S. Haswell, R. Phillips, D.C. Rees, *Structure* **19**, 1356 (2011)
- K. Hayakawa, H. Tatsumi, M. Sokabe, *J. Cell Sci.* **121**, 496 (2008)
- P.S. Heckbert, M. Garland, *Comp. Geom. Theor. Appl.* **14**, 49 (1999)
- B. Honig, A. Nicholls, *Science* **268**, 1144 (1995)
- W. Im, D. Beglov, B. Roux, *Comput. Phys. Commun.* **111**, 59 (1998)
- W. Im, M. Feig, C.L. Brooks, *Biophys. J.* **85**, 2900 (2003)
- D.E. Ingber, *FASEB J.* **20**, 811 (2006)
- H.I. Ingolfsson, C.A. Lopez, J.J. Uusitalo, D.H. de Jong, S.M. Gopal, X. Periole, S.J. Marrink, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **4**, 225 (2014)
- I. Iscla, P. Blount, *Biophys. J.* **103**, 169 (2012)
- I. Iscla, R. Wray, P. Blount, *Biophys. J.* **95**, 2283 (2008)
- J. Jeon, G.A. Voth, *Biophys. J.* **94**, 3497 (2008)
- K. Kamaraju, V. Belyy, I. Rowe, A. Anishkin, S. Sukharev, *J. Gen. Physiol.* **138**, 49 (2011)
- M. Karplus, J. Kuriyan, *Proc. Natl. Acad. Sci. U. S. A.* **102**, 6679 (2005)
- J.L. Klepeis, K. Lindorff-Larsen, R.O. Dror, D.E. Shaw, *Curr. Opin. Struct. Biol.* **19**, 120 (2009)
- A. Kloda, B. Martinac, *EMBO J.* **20**, 1888 (2001)
- P. Koprowski, A. Kubalski, *J. Membr. Biol.* **164**, 253 (1998)
- P. Koprowski, W. Grajkowski, E.Y. Isacoff, A. Kubalski, *J. Biol. Chem.* **286**, 877 (2011)
- O. Krishtal, *Trends Neurosci.* **26**, 477 (2003)
- J.Y. Lai, Y.S. Poon, J.T. Kaiser, D.C. Rees, *Protein science: A publication of the protein. Society* **22**, 502 (2013)
- G. Levin, P. Blount, *Biophys. J.* **86**, 2862 (2004)
- N. Levina, S. Totemeyer, N.R. Stokes, P. Louis, M.A. Jones, I.R. Booth, *EMBO J.* **18**, 1730 (1999)
- M. Louhivuori, H.J. Risselada, E. van der Giessen, S.J. Marrink, *Proc. Natl. Acad. Sci. U. S. A.* **107**, 19856 (2010)
- L. Ma, A. Yethiraj, X. Chen, Q. Cui, *Biophys. J.* **96**, 3543 (2009)
- H. Machiyama, H. Tatsumi, M. Sokabe, *Biophys. J.* **97**, 1048 (2009)
- H.R. Malcolm, Y.Y. Heo, D.E. Elmore, J.A. Maurer, *Biophys. J.* **101**, 345 (2011)
- V.S. Markin, F. Sachs, *Phys. Biol.* **1**, 110 (2004)
- S.J. Marrink, D.P. Tieleman, *Chem. Soc. Rev.* **42**, 6801 (2013)
- S.J. Marrink, H.J. Risselada, S. Yefimov, D.P. Tieleman, A.H. de Vries, *J. Phys. Chem. B* **111**, 7812 (2007)
- B. Martinac, *J. Cell Sci.* **117**, 2449 (2004)
- B. Martinac, A. Kloda, *Prog. Biophys. Mol. Biol.* **82**, 11 (2003)

- B. Martinac, A. Kloda, *Comprehensive Biophysics* (Elsevier, Amsterdam, 2012), p. 108
- B. Martinac, M. Buechner, A.H. Delcour, J. Adler, C. Kung, Proc. Natl. Acad. Sci. U. S. A. **84**, 2297 (1987)
- B. Martinac et al., *Antioxid. Redox Signal.* **20**, 952 (2014)
- M. Masetti, C. Berti, R. Ocello, G.P. Di Martino, M. Recanatini, C. Fiegna, A. Cavalli, J. Chem. Theory Comput. **12**, 5681 (2016)
- S. Miller, W. Bartlett, S. Chandrasekaran, S. Simpson, M. Edwards, I.R. Booth, EMBO J. **22**, 36 (2003)
- P. Moe, P. Blount, *Biochemistry* **44**, 12239 (2005)
- L. Monticelli, S.K. Kandasamy, X. Periole, R.G. Larson, D.P. Tieleman, S.J. Marrink, J. Chem. Theory Comput. **4**, 819 (2008)
- T. Nomura, M. Sokabe, K. Yoshimura, *Biophys. J.* **91**, 2874 (2006)
- T. Nomura, M. Sokabe, K. Yoshimura, *Biophys. J.* **94**, 1638 (2008)
- T. Nomura, M. Sokabe, K. Yoshimura, *BioMed Res. Int.* **2016**, 2401657 (2016)
- A. Pantano, M.C. Boyce, D.M. Parks, *Phys. Rev. Lett.* **91**, 145504 (2003)
- E. Perozo, D.M. Cortes, P. Sompornpisut, A. Kloda, B. Martinac, *Nature* **418**, 942 (2002a)
- E. Perozo, A. Kloda, D.M. Cortes, B. Martinac, *Nat. Struct. Biol.* **9**, 696 (2002b)
- E. Petrov, D. Palanivelu, M. Constantine, P.R. Rohde, C.D. Cox, T. Nomura, D.L. Minor Jr., B. Martinac, *Biophys. J.* **104**, 1426 (2013)
- R. Phillips, T. Ursell, P. Wiggins, P. Sens, *Nature* **459**, 379 (2009)
- C. Pliotas, J.H. Naismith, *Curr. Opin. Struct. Biol.* **45**, 59 (2016)
- C. Pliotas et al., Proc. Natl. Acad. Sci. U. S. A. **109**, E2675 (2012)
- C. Pliotas et al., *Nat. Struct. Mol. Biol.* **22**, 991 (2015)
- A.M. Powl, J.M. East, A.G. Lee, *Biochemistry* **42**, 14306 (2003)
- A.M. Powl, J.M. East, A.G. Lee, *Biochemistry* **44**, 5873 (2005a)
- A.M. Powl, J.N. Wright, J.M. East, A.G. Lee, *Biochemistry* **44**, 5713 (2005b)
- A.M. Powl, J.M. East, A.G. Lee, *Biochemistry* **47**, 12175 (2008)
- M. Praprotnik, L.D. Site, K. Kremer, *Annu. Rev. Phys. Chem.* **59**, 545 (2008)
- B.L. Pruitt, A.R. Dunn, W.I. Weis, W.J. Nelson, *PLoS Biol.* **12**, e1001996 (2014)
- I. Rowe, A. Anishkin, K. Kamaraju, K. Yoshimura, S. Sukharev, *J. Gen. Physiol.* **143**, 543 (2014)
- H. Sackin, *Annu. Rev. Physiol.* **57**, 333 (1995)
- Y. Saimi, B. Martinac, A.H. Delcour, P.V. Minorsky, M.C. Gustin, M.R. Culbertson, J. Adler, C. Kung, *Methods Enzymol.* **207**, 681 (1992)
- M.F. Sanner, A.J. Olson, J.C. Spohner, *Biopolymers* **38**, 305 (1996)
- M.G. Saunders, G.A. Voth, *Curr. Opin. Struct. Biol.* **22**, 144 (2012)
- Y. Sawada, M. Murase, M. Sokabe, *Channels (Austin)* **6**, 317 (2012)
- F. Scarpa, S. Adhikari, A.J. Gil, C. Remillat, *Nanotechnology* **21**, 125702 (2010)
- U. Schumann, M.D. Edwards, C. Li, I.R. Booth, *FEBS Lett.* **572**, 233 (2004)
- Q. Shi, S. Izvekova, G.A. Voth, *J. Phys. Chem. B* **110**, 15045 (2006)
- W. Shinoda, R. DeVane, M.L. Klein, *Curr. Opin. Struct. Biol.* **22**, 175 (2012)
- C.D. Snow, E.J. Sorin, Y.M. Rhee, V.S. Pande, *Annu. Rev. Biophys. Biomol. Struct.* **34**, 43 (2005)
- M. Sotomayor, K. Schulten, *Biophys. J.* **87**, 3050 (2004)
- M. Sotomayor, T.A.v.d. Straaten, U. Ravaioli, K. Schulten, *Biophys. J.* **90**, 3496 (2006)
- S.A. Spronk, D.E. Eldmore, D.A. Dougherty, *Biophys. J.* **90**, 3555 (2006)
- S. Steinbacher, R. Bass, P. Strop, D.C. Rees, *Mechanosens. Ion Channels Part A* **58**, 1 (2007)
- S. Sukharev, *Biophys. J.* **83**, 290 (2002)
- S. Sukharev, A. Anishkin, *Trends Neurosci.* **27**, 345 (2004)
- S. Sukharev, D.P. Corey, *Sci. Signal.* **2004**, re4 (2004)
- S.I. Sukharev, P. Blount, B. Martinac, C. Kung, *Annu. Rev. Physiol.* **59**, 633 (1997)
- S.I. Sukharev, W.J. Sigurdson, C. Kung, F. Sachs, *J. Gen. Physiol.* **113**, 525 (1999)
- S. Sukharev, M. Betanzos, C.S. Chiang, H.R. Guy, *Nature* **409**, 720 (2001a)
- S. Sukharev, S.R. Durell, H.R. Guy, *Biophys. J.* **81**, 917 (2001b)
- H. Sun, D.P. Li, S.R. Chen, W.N. Hittelman, H.L. Pan, *J. Pharmacol. Exp. Ther.* **331**, 851 (2009)
- Y. Tang, G. Cao, X. Chen, J. Yoo, A. Yethiraj, Q. Cui, *Biophys. J.* **91**, 1248 (2006)

- Y. Tang, J. Yoo, A. Yethiraj, Q. Cui, X. Chen, *Biophys. J.* **95**, 581 (2008)
- A. Torres-Sánchez, J.M. Vanegas, M. Arroyo, *Phys. Rev. Lett.* **114**, 258102 (2015)
- I.J. Tsai, Z.W. Liu, J. Rayment, C. Norman, A. McKinley, B. Martinac, *Eur. Biophys. J.* **34**, 403 (2005)
- K.I. Tserpes, P. Papanikos, *Compos. Struct.* **91**, 131 (2009)
- M.S. Turner, P. Sens, *Phys. Rev. Lett.* **93**, 118103 (2004)
- D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J. Berendsen, *J. Comput. Chem.* **26**, 1701 (2005)
- J.M. Vanegas, M. Arroyo, *PLoS One* **9**, e113947 (2014)
- J.M. Vanegas, A. Torres-Sánchez, M. Arroyo, *J. Chem. Theory Comput.* **10**, 691 (2014)
- V. Vasquez, *Biophys. J.* **104**, 1391 (2013)
- V. Vasquez, M. Sotomayor, J. Cordero-Morales, K. Schulten, E. Perozo, *Science* **321**, 1210 (2008a)
- V. Vasquez, M. Sotomayor, D.M. Cortes, B. Roux, K. Schulten, E. Perozo, *J. Mol. Biol.* **378**, 55 (2008b)
- R.M. Venable, F.L. Brown, R.W. Pastor, *Chem. Phys. Lipids* **192**, 60 (2015)
- T. Vora, B. Corry, S.H. Chung, *Biochim. Biophys. Acta* **1758**, 730 (2006)
- J.A. Wagoner, N.A. Baker, *Proc. Natl. Acad. Sci.* **103**, 8331 (2006)
- H.C. Wang, B.P. Thampatty, *Biomech. Model. Mechanobiol.* **5**, 1 (2006)
- W. Wang, S.S. Black, M.D. Edwards, S. Miller, E.L. Morrison, W. Bartlett, C. Dong, J.H. Naismith, I.R. Booth, *Science* **321**, 1179 (2008)
- Y. Wang, Y. Liu, H.A. Deberg, T. Nomura, M.T. Hoffman, P.R. Rohde, K. Schulten, B. Martinac, P.R. Selvin, *elife* **3**, e01834 (2014)
- R. Ward et al., *Biophys. J.* **106**, 834 (2014)
- P. Wiggins, R. Phillips, *Proc. Natl. Acad. Sci. U. S. A.* **101**, 4071 (2004)
- K.M. Wisdom, S.L. Delp, E. Kuhl, *Biomech. Model. Mechanobiol.* **14**, 195 (2014)
- S. Yefimov, E. van der Giessen, P.R. Onck, S.J. Marrink, *Biophys. J.* **94**, 2994 (2008)
- S.O. Yesylevskyy, L.V. Schafer, D. Sengupta, S.J. Marrink, *PLoS Comput. Biol.* **6**, e1000810 (2010)
- Y. Zeng, A.K. Yip, S.K. Teo, K.H. Chiam, *Biomech. Model. Mechanobiol.* **11**, 49 (2012)
- L. Zhu, J. Wu, L. Liu, Y. Liu, Y. Yan, Q. Cui, X. Chen, *Biomech. Model. Mechanobiol.* **15**, 1557 (2016)



Spherical Indentation on a Prestressed Elastic Coating/Substrate System

4

James A. Mills and Xi Chen

Contents

Introduction	130
Model and Computation Method	132
Formulation for a Fixed Indenter Radius	134
Forward Analysis	134
Reverse Analysis	135
General Formulation with Variable Indenter Radius	140
Forward Analysis	140
Reverse Analysis	141
Error Sensitivity	141
Conclusion	150
References	151

Abstract

While there have been many studies on the indentation test of thin film/substrate systems, the primary goal has been determining the film properties. However, there was very little effort to probe the properties of both the film and the substrate (the latter may be as important as the film properties). Moreover, a prestress usually exists in the film, typically resulted from mismatched deformation or material properties. In this study, we establish a spherical indentation framework to examine the material properties of both the film and substrate as well as

J. A. Mills

Department of Civil Engineering and Engineering Mechanics, Columbia University, New York, NY, USA

X. Chen (✉)

Department of Earth and Environmental Engineering, Columbia Nanomechanics Research Center, Columbia University, New York, NY, USA

e-mail: xichen@columbia.edu

determining the film prestress. An indentation test is performed at two prescribed depths, and functional forms are established between indentation parameters and material variables. An effective reverse analysis algorithm is established to deduce the desired material properties. The potential error sensitivity is also examined in a systematic way. The study on prestressed film/substrate systems has many potential applications in engineering.

Keywords

Indentation · Thin film · Substrate · Elastic property · Film prestress

Introduction

Film or substrate composites are commonly found in the building, appliance, electronics, and transportation industry and are considered well-engineered systems. More specifically in building structures, the film/substrate composites exist in the cladding of a roof or siding sandwich panel, and an optically coated glazing system acts as a part of the building façade. The film/substrates are fundamental building blocks of electronic devices, where typically one or more layers of metallic or semiconductor films are deposited onto a substrate material, most notably silicon. Today's optical devices usually have some kind of coating deposited on the lens: possibly a scratch-resistant coating, an antireflective coating, an ultraviolet treatment, or a combination of all three. The skin/tissue also acts as a film/substrate system with the prestressed skin acting as a protective and sensing layer. The paint applied to an automobile not only helps prevent corrosion but also adds to the visual appearance. Turbine blades are the prime movers for jet aircraft and oftentimes have a heat-resistant coating applied to its surface. In each of the specific examples given, the film must stay fully attached to the substrate.

While the mechanical properties of the film and substrate usually differ, they act together as a complete structural system. Knowing the mechanical properties (e.g., modulus, yield strength) of each is a requisite for fully understanding the response of the system or the level of deterioration that may have occurred over time to the system. Measuring the film thickness is also quite useful as it may give us a good indication as to the level of deterioration of the system or the continued effectiveness of the system. For example, measuring the thickness of a turbine blade's thermal coating may be a good indicator of continued safe operation of the turbine. However, this oftentimes requires the cutting of the blade and exposing the cross section. To have the ability to determine the coating thickness without having to cut the blade or possibly not even having to remove the blade from the turbine assembly certainly would be less costly and more desirable.

One of the exciting and relatively inexpensive techniques for measuring the mechanical properties of a film/substrate system is the use of micro- or nanoindentation equipment. It is arguably the simplest approach to measuring the properties of small-scale structural assemblies including thin films (Chen and Vlassak 2001; Pharr 1998). The biggest advantage in using this method is that it does not require

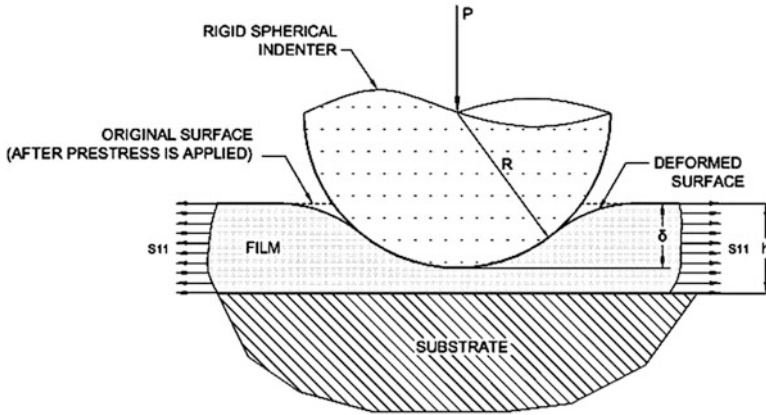


Fig. 1 Indenter/film/substrate model

the film to be separated from the substrate. An indentation test is performed utilizing a hardened (diamond) tipped indenter, being pushed into a film/substrate system (Fig. 1).

During this process, the force required (indentation load, P) to push the indenter into the material is measured simultaneously with the indentation depth (δ). While the stress fields generated from the nonlinear contact and large deformation during the indentation test are quite complicated, it has been found that the P - δ data can implicitly relate the film elastoplastic properties (Gao et al. 1992; Hu and Lawn 1998). Using this P - δ data along with a functional relationship established in the forward analysis, one can predict the material/mechanical properties through a reverse analysis (Cheng and Cheng 2004).

The present interest is not only to further understand the mechanical properties of the film but also the properties of the substrate, along with examining these properties under an equibiaxial film prestress state. The prestress may arise from thermal expansion mismatch, lattice mismatch, peening, etc., and it can be quite high sometimes and responsible for premature failure or delamination of the film from the substrate (Hu et al. 1988; Evans and Hutchinson 1984). Suresh and Giannakopoulos (1998) proposed the use of sharp indenters to measure the residual stress of thin films and bulk materials. However this method requires multiple indentation tests. In particular it requires a reference specimen with zero prestress whose indentation behavior is compared with the specimen where prestress is present. This sometimes is impractical or difficult to achieve.

Another consideration with regard to indentation testing of film/substrate systems has to do with the often overlooked determination of substrate mechanical properties. Conventional approaches focus on probing the film material properties (Cheng and Cheng 2004; Doerner and Nix 1986; Gouldstone et al. 2007; Saha and Nix 2002), and in order to avoid the substrate effect, the indentation depth was typically kept below 10–20% of the film thickness (h) (Tsui et al. 1999; Cheng and

Cheng 2004). On one hand, there are inherent problems associated with performing indentation tests at shallow depths (Saha et al. 2001; Swadener et al. 2002; Chen and Vlassak 2001), that at times makes the test unattractive. For example, the P - δ data of shallow indentation may be affected by localized material surface roughness, indenter tip shape changes due to wear, or possible indenter drag (frictional effects) from surface adhesion. On the other hand, there are instances where the substrate properties are of interest as well. The turbine blade noted earlier is a fine example of where the blade properties (substrate) may be as important as the thermal coating.

The aforementioned gaps may be bridged by performing a moderately deep test, embracing and fully incorporating in the present analysis the coupled effect from both the film and substrate (Chen and Vlassak 2001; Chen et al. 2006; Zhao et al. 2006, 2007), as well as accounting for the film prestress. A relationship accounting for both the film and substrate effects in the P - δ data can be developed based on which an effective reverse analysis algorithm can be established. The moderately deep indentation measurement may help to yield a unique solution for determining the material property through reverse analysis (Chen et al. 2007).

The goal of the present study is to establish the framework of indenting a prestressed film on a semi-infinite substrate, using only one spherical indenter (which prevents penetrating/damaging the film) and without requiring a stress-free reference specimen. The focus is given to the elastic properties of the film and substrate, specifically determining the elastic modulus of each, as well as the film prestress. The next section describes the physical model as well as the parameter groups. Based on extensive finite element simulations of indentation tests over a wide range of material parameters and prestress levels, a set of general functional relationships are developed through the forward analysis. Through the results of two tests at different depths, the reverse analysis can effectively identify two material parameters among the film elastic modulus, substrate elastic modulus, and film prestress (as long as the third one is known a priori). The study is further extended to variable indenter radii or variable film thickness. Finally, a rigorous review of the error sensitivity in the proposed model is discussed.

Model and Computation Method

As shown in Fig. 1, the model is made up of a rigid spherical indenter, and a thin film is fully affixed to a semi-infinite substrate. Both the film and the substrate are modeled as homogeneous, isotropic elastic materials. The film prestress is equibiaxial, σ_{ps} . The indenter acts directly against the film and has a prescribed displacement that is normal to the upper surface of the film. The film has an elastic modulus defined as E_f and Poisson's ratio of ν_f . The substrate has an elastic modulus of E_s and a corresponding Poisson's ratio, ν_s . ν_s and ν_f are fixed as 0.25, and they are minor factors for indentation analysis (Cheng and Cheng 1998).

The study is primarily focused on, through a single indentation test, the determination of two of three variables: σ_{ps} , E_f , and E_s (as long as the third is known a

priori). Invoking dimensional analysis, the following functional relationship needs to be determined through forward analysis:

$$\frac{P}{E_f \delta^2} = \Pi \left(\frac{E_f}{E_s}, \frac{R}{h}, \frac{\sigma_{ps}}{E_f}, \frac{\delta}{h} \right) \quad (1)$$

To take advantage of the substrate effect (Zhao et al. 2006, 2007; Chen et al. 2006; Zhao et al. 2006), the indentation load can be taken at two indentation depths $\delta_1 = h/2$, and $\delta_2 = h/4$, where h is the film thickness (Mills and Chen 2009), from which two independent relationships are deduced:

$$\frac{P_1}{E_f \delta_1^2} = f_1 \left(\frac{E_f}{E_s}, \frac{R}{h}, \frac{\sigma_{ps}}{E_f} \right) \quad (2)$$

$$\frac{P_2}{E_f \delta_2^2} = f_2 \left(\frac{E_f}{E_s}, \frac{R}{h}, \frac{\sigma_{ps}}{E_f} \right) \quad (3)$$

Subscripts 1 and 2 relate the functional relationships in (2) and (3) to the two prescribed indentation depths: δ_1 and δ_2 . Note that in practice, the maximum indentation depth can exceed $h/2$ (and be any value), and only the data taken at $h/2$ and $h/4$ are relevant.

As a first step, it is assumed that the film thickness (h) is known such that one may correspondingly match the indenter radius (R) exactly to this film thickness (R/h will be varied in a later part of this paper). In this simplified case:

$$\frac{P_1}{E_f \delta_1^2} = f_1 \left(\frac{E_f}{E_s}, \frac{\sigma_{ps}}{E_f} \right), \quad \text{with } \frac{R}{h} = 1.0 \quad (4)$$

$$\frac{P_2}{E_f \delta_2^2} = f_2 \left(\frac{E_f}{E_s}, \frac{\sigma_{ps}}{E_f} \right), \quad \text{with } \frac{R}{h} = 1.0 \quad (5)$$

After these two equations are established through forward analysis, one will be able to solve for two unknowns E_f , E_s , or σ_{ps} with the third being known a priori.

Numerical simulation of indentation was performed using ABAQUS (Dassault Systèmes (SIMULIA) 2008). The film/substrate was made up of axisymmetric 4-node bilinear, reduced integration elements (CAX4R). Over 180,000 elements (Fig. 2) were incorporated into the model, and the substrate is semi-infinite. The film is equibiaxially prestressed to a desired level prior to indentation. The contact is assumed frictionless which is a relatively minor factor (Bucaille et al. 2003; Cheng and Cheng 2004).

The material properties are varied in moderate ranges during the forward analysis, with the elastic modulus ratio ($EMR = \frac{E_f}{E_s}$) in the range of $0.25 \leq EMR \leq 15.00$ and the normalized prestress ($K = \sigma_{ps}/E_f$) between $-0.10 \leq K \leq 0.10$. It is verified that film buckling does not take place under such a prestress level.

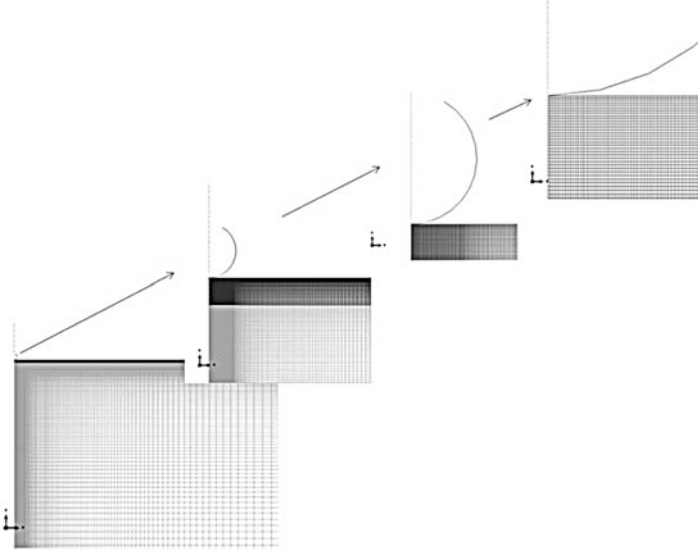


Fig. 2 Finite element model

Formulation for a Fixed Indenter Radius

Forward Analysis

As can be seen in the normalized indentation load-depth relationships in Fig. 3, the substrate effect as well as the spherical geometry indenter plays a significant role in creating a nonlinear force-displacement relationship (Chen et al. 2007). In comparing the curvatures for $EMR = 0.25$ in Fig. 4a and $EMR = 15.0$ in Fig. 4b, as the substrate becomes more flexible (less stiff), the plots take on a more linear formulation, which is due to the loss in substrate effect and the indentation behavior becomes dominated by the film-bending effect.

The presence of prestress also affects the indentation force, and the effect is more prominent for the more compliant substrate. As the film prestress increases, the normalized indentation force also increases so as to overcome the tension in film. Through the fitting of extensive numerical simulations, for $\delta_1 = h/2$ and $\delta_2 = h/4$, a set of general functions was formulated according to Eqs. (4) and (5):

$$\frac{P_i}{E_f \delta_i^2} = f_i \left(\frac{E_f}{E_s}, \frac{\sigma_{ps}}{E_f} \right) = A_i \left(\frac{E_f}{E_s} - B_i \right)^{C_i} + D_i \left(\frac{E_f}{E_s} \right)^2 + E_i \left(\frac{E_f}{E_s} \right) + F_i \quad (6)$$

$$i = 1, 2 \text{ for } \delta_1 = \frac{h}{2} \text{ and } \delta_2 = \frac{h}{4}$$

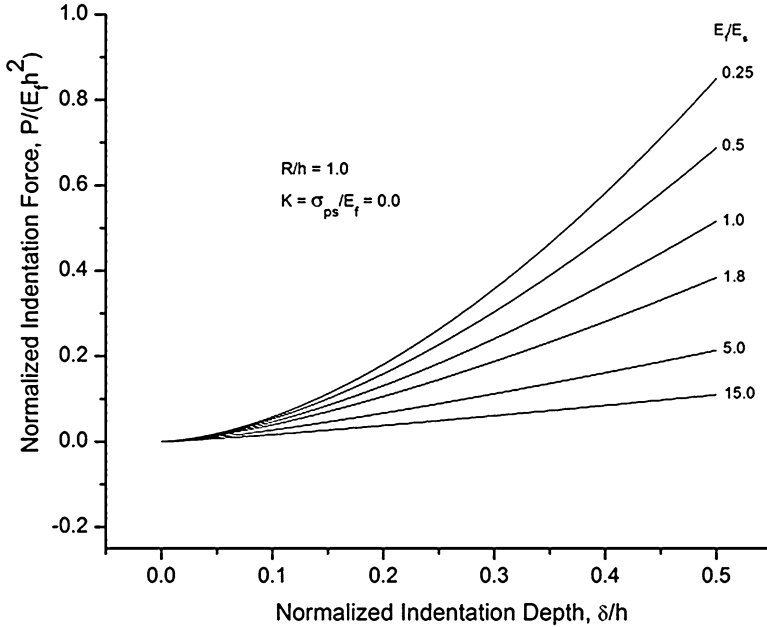


Fig. 3 Normalized indentation load – penetration depth for various elastic modulus ratios (EMR), with $R/h = 1.0$, $K = 0$

with

$$\begin{aligned}
 A_i &= a_{1i} K + a_{2i} K^2 + a_{3i} K^3 + a_{4i} K^4 \\
 B_i &= b_{1i} K + b_{2i} K^2 + b_{3i} K^3 + b_{4i} K^4 \\
 C_i &= C_{1i} (c_{2i} - K)^{C_{3i}} \\
 D_i &= d_{1i} K + d_{2i} K^2 + d_{3i} K^3 + d_{4i} K^4 \\
 E_i &= e_{1i} K + e_{2i} K^2 + e_{3i} K^3 + e_{4i} K^4 \\
 F_i &= \text{constant}
 \end{aligned}$$

where coefficients a_{ji} , b_{ji} , c_{ki} , d_{ji} , e_{ji} ($i = 1-2$, $j = 1-4$, $k = 1-3$) are fitted constants as shown in Table 1. The representative Eq. (6) is also shown as surface plots in Fig. 5a, b, which correspond to prescribed indenter displacements of δ_1 and δ_2 , respectively. The data points generated from the forward analysis (via numerical simulation) are also shown, with excellent agreement with the surface plots.

Reverse Analysis

The two independent equations in (6) allow the solution of any two of the three unknowns E_f , E_s , or σ_{ps} , as long as the third term is known accurately. The reverse

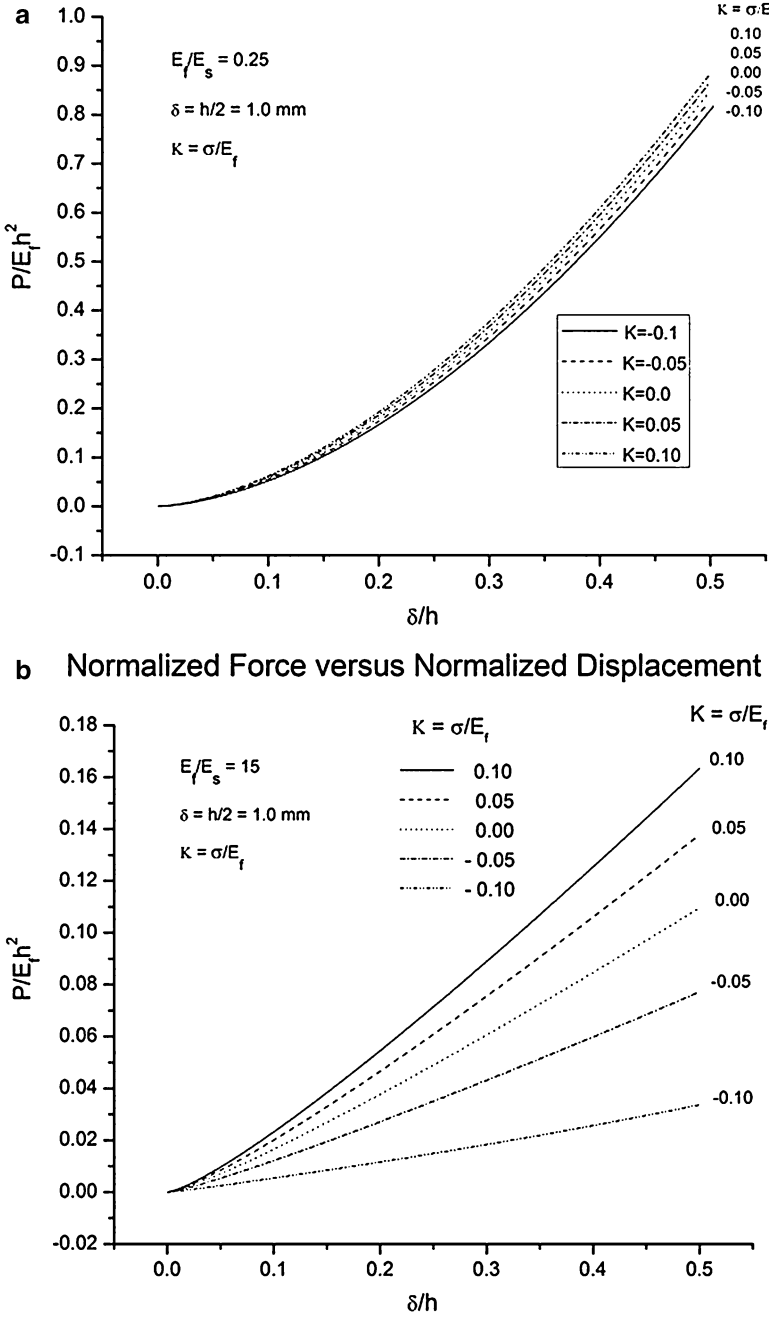


Fig. 4 Normalized load versus normalized indentation depth for various normalized prestress values at (a) $E_f/E_s = 0.25$ and (b) $E_f/E_s = 15.0$

Table 1 General coefficients for EMR-K fitting function

F_1	-0.138142626	F_2	-0.27345413
a_{01}	2.674281604	a_{02}	4.034252118
a_{11}	0.802831035	a_{12}	1.147974895
a_{21}	1.971999573	a_{22}	5.222286008
a_{31}	-8.544787918	a_{32}	-29.25982734
a_{41}	44.46597133	a_{42}	131.1351472
b_{01}	-0.383265786	b_{02}	-0.568865038
b_{11}	0.549111802	b_{12}	0.763772049
b_{21}	-1.29341713	b_{22}	-2.600832556
b_{31}	3.373917099	b_{32}	10.53364711
b_{41}	-24.85527757	b_{42}	-60.95042138
c_{11}	-0.35865278	c_{12}	-0.29205364
c_{21}	-0.434558478	c_{22}	-0.295299789
c_{31}	-0.63882907	c_{32}	-0.470757762
d_{01}	-0.000225813	d_{02}	-0.000201266
d_{11}	-0.003103271	d_{12}	-0.003870197
d_{21}	0.003580859	d_{22}	-0.00153943
d_{31}	0.102776788	d_{32}	0.277177385
d_{41}	-0.877371942	d_{42}	-2.443837931
e_{01}	0.008209149	e_{02}	0.008712058
e_{11}	0.110259779	e_{12}	0.157132416
e_{21}	-0.24475966	e_{22}	-0.276338029
e_{31}	-0.115013816	e_{32}	-1.035169875
e_{41}	3.664067224	e_{42}	13.69568226

analysis is based on minimizing the total error of the two equations (Mills and Chen 2009), utilizing P - δ data from the corresponding numerical indentation experiment (assuming h is known and $R/h = 1$ is used in the experiment). To verify the reverse analysis procedure, a number of separate numerical indentation tests were performed, using EMR and K parameters that were independent from the parameters used in the forward analysis.

As a first example, assuming E_f is known, the error associated with E_s and σ_{ps} deduced from reverse analysis (with respect to that used in forward analysis) is shown in Fig. 6a, with most errors less than 3% and maximum error about 7%. In the second example (Fig. 6b), with E_s known, the error associated with the reverse analysis of E_f and σ_{ps} is mostly below 2%, and the maximum error is about 6%. Finally, if the prestress is known in advance, the accuracy for determining the film and substrate moduli is very good, with errors well below 1% in most cases (Fig. 6c).

A close examination shows that the maximum errors tend to occur at the edges of the defined problem parameters. That is, the relatively more prominent errors (though only about 7%) are likely to occur at high compressive prestress and at a correspondingly high film moduli (recall that the parameter ranges investigated

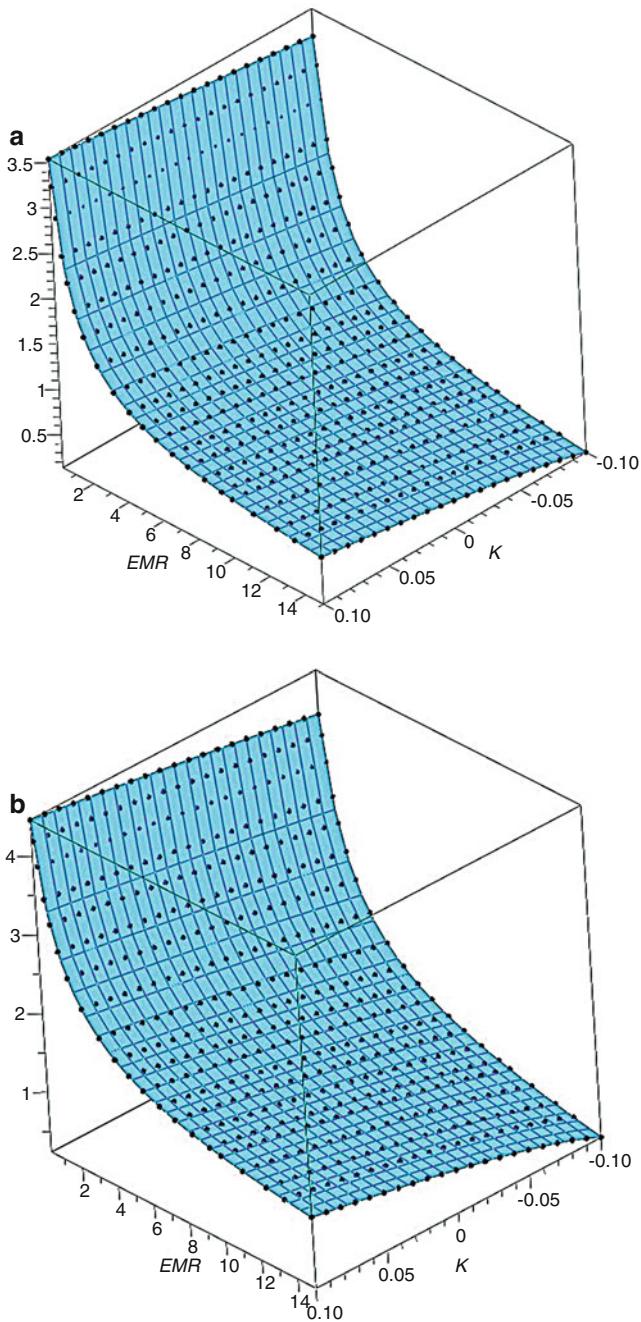


Fig. 5 Fitting function surface plot with forward analysis data points at (a) $\delta_1 = h/2$ and (b) $\delta_2 = h/4$

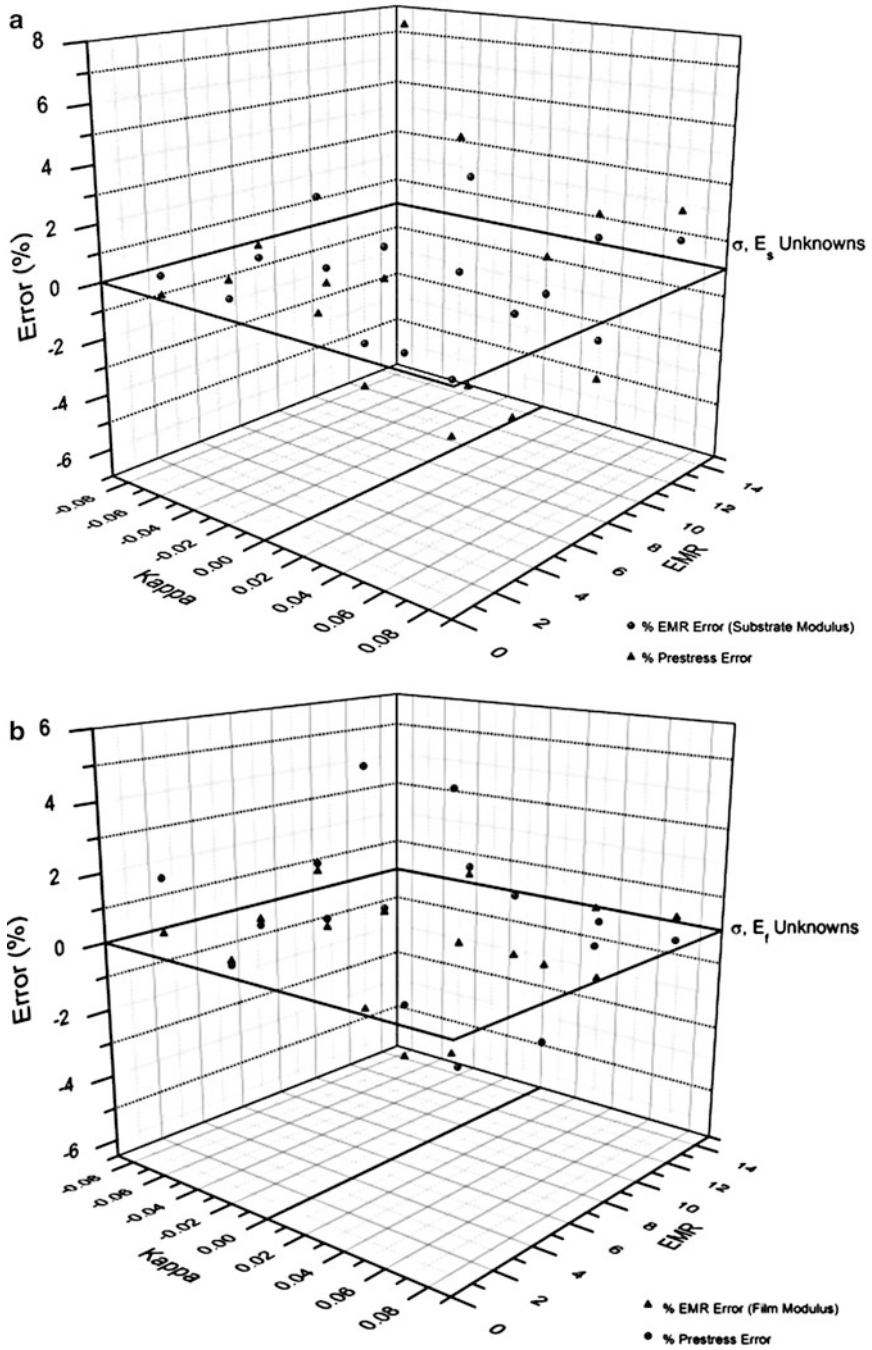


Fig. 6 (continued)

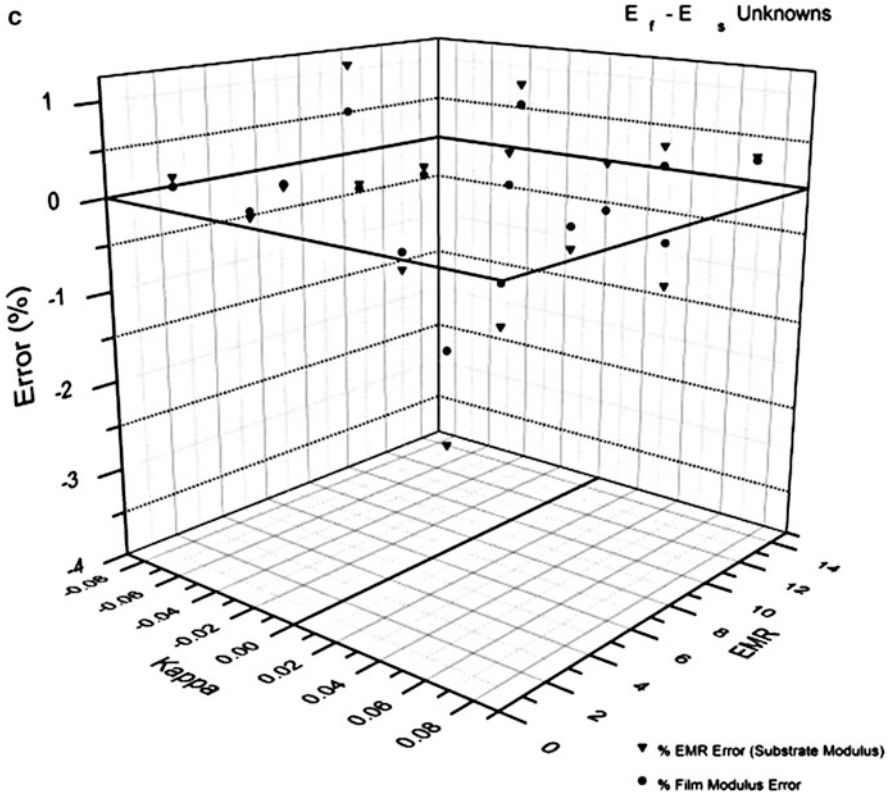


Fig. 6 Reverse analysis error results (a) with σ and E_s as unknowns; (b) with σ and E_f as unknowns; (c) with E_s and E_f as unknowns

in this study are $0.25 \leq \frac{E_f}{E_s} \leq 15.00$ and $-0.10 \leq K \leq 0.10$). This may be due to the inherent instability as the normalized prestress approaches -0.10 . Overall, the reverse analysis algorithm is fairly robust and is valid for moderately large ranges of film and substrate moduli and film prestress.

General Formulation with Variable Indenter Radius

Forward Analysis

In many instances, the film thickness may not be known or may not be able to be measured. Under certain circumstances, it would be useful to extend the functional relationships developed in section “[Formulation for a Fixed Indenter Radius](#)” to the problem of unknown film thickness (or conversely to remove the testing

constraint of having to match the indenter radius to the film thickness). In order to show the feasibility without complicating the formulation too much, a proof-of-principle investigation is conducted in this section by choosing a fixed EMR value ($EMR = 2.5$) whereas allowing K and R/h to vary as $-0.10 \leq K \leq 0.10$ and R/h of 0.5, 0.75, 1.0, 2.0, 5.0, and 10.0, respectively. Plots of the normalized load versus prestress level are given in Fig. 7a, b for the two indentation depths. Each plot shows various sets of curves that were generated at different R/h values, and the separation between them provides confidence for deducing R/h through reverse analysis.

Through extensive forward analyses, the following equations are established:

$$\frac{P_i}{E_f \delta_i^2} = M_i + N_i K + Q_i K^2 \quad \text{where } i = 1, 2 \quad \text{and } \delta_1 = h/2 \text{ and } \delta_2 = h/4 \quad (7)$$

$$\begin{aligned} M_i &= m_{1i} e^{[(R/h)/t_{1i}]} + m_{2i} e^{[(R/h)/t_{2i}]} + y_{0i} \\ N_i &= n_{1i} e^{[(R/h)/t_{3i}]} + n_{2i} e^{[(R/h)/t_{4i}]} + y_{1i} \\ Q_i &= q_{1i} + q_{2i} (R/h) + q_{3i} q_{4i}^{(R/h)} \end{aligned}$$

Coefficients in M_i , N_i , and Q_i are shown in Table 2.

Reverse Analysis

By performing a single indentation test, the solution of two variables, σ_{ps} and h , can be obtained. In Fig. 8, R/h values were chosen to be 0.6, 4.0, and 9.0 with corresponding σ_{ps}/E_f values of -0.1 , 0 , and 0.1 . Through the minimization of error of Eq. (7), the reverse analysis result errors (deduced parameter vs. the input parameter) are mostly below 1–2%, with the largest error of about 6%. Note that since h is an unknown parameter to be determined, the reverse analysis algorithm, Fig. 9, is an iterative, averaging technique with the only requirement being that the indentation experiment (FEA) be performed to a depth at least as deep as $\delta = h/2$.

Note that (7) is specified for $EMR = 2.5$; nevertheless, the approach can be readily extended to other EMR values and that would add another dimension in reverse analysis, making the formulation more complicated, yet the strategy remains straightforward.

Error Sensitivity

When performing an indentation study, the accuracy of the P - δ data (P_1 , P_2 , δ_1 , δ_2) is assumed to be exact. In practice, measurement errors can affect the accuracy of the deduced material properties. For these reasons, the error sensitivity

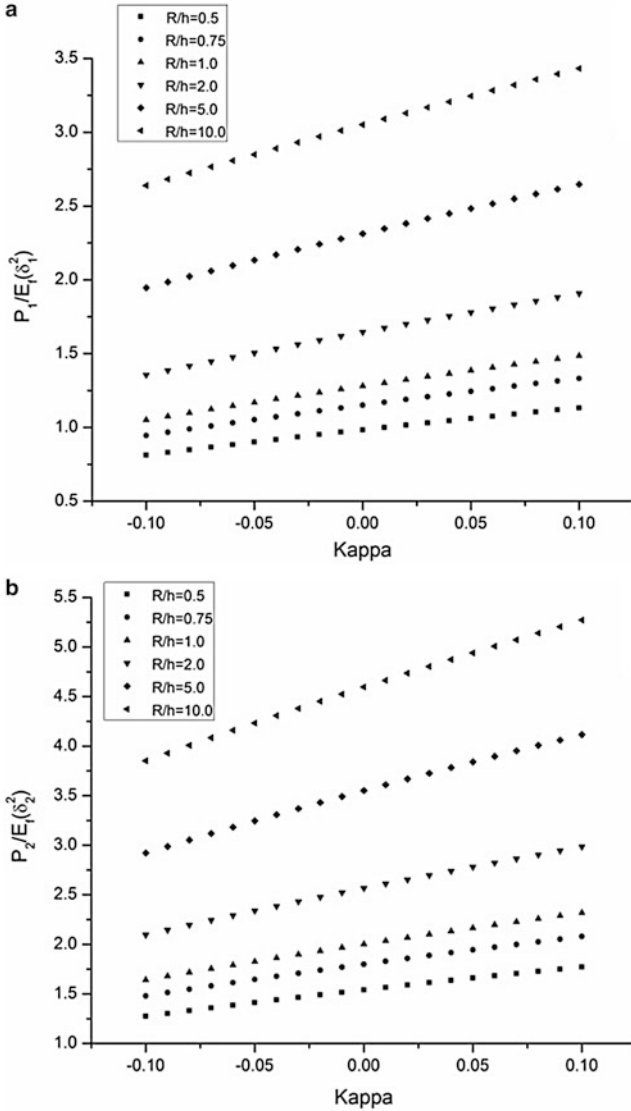


Fig. 7 Normalized indentation load versus $K = \sigma_{ps}/E_f$, at $0.5 \leq R/h \leq 10.0$, (a) $\delta_1 = h/2$. (b) $\delta_2 = h/4$

and its significance to each variable in the fitting function should be examined. Note that in this analysis the film thickness is assumed to be known, and the indenter radius is matched accordingly.

Differentiating the dimensionless functions from (4) and (5) term by term, we get the following:

Table 2 Coefficients for R/h-K fitting function

M_i			
Y_{01}	4.292937	Y_{02}	6.174527
m_{11}	-3.15663	m_{21}	-4.35566
m_{21}	-0.64938	m_{22}	-1.03513
t_{11}	-10.7228	t_{12}	-9.8578
t_{21}	-0.63728	t_{22}	-0.6755
N_i			
Y_{11}	4.163368	Y_{12}	7.747632
n_{11}	-1.4351	n_{12}	-2.01002
n_{21}	-2.19215	n_{22}	-4.76786
t_{31}	-0.59133	t_{32}	-0.66103
t_{41}	-4.20076	t_{42}	-5.08646
Q_i			
q_{11}	-1.70805	q_{12}	-3.39752
q_{21}	0.018993	q_{22}	-6.92E-04
q_{31}	0.807144	q_{32}	2.102068
q_{41}	0.475559	q_{42}	0.582964

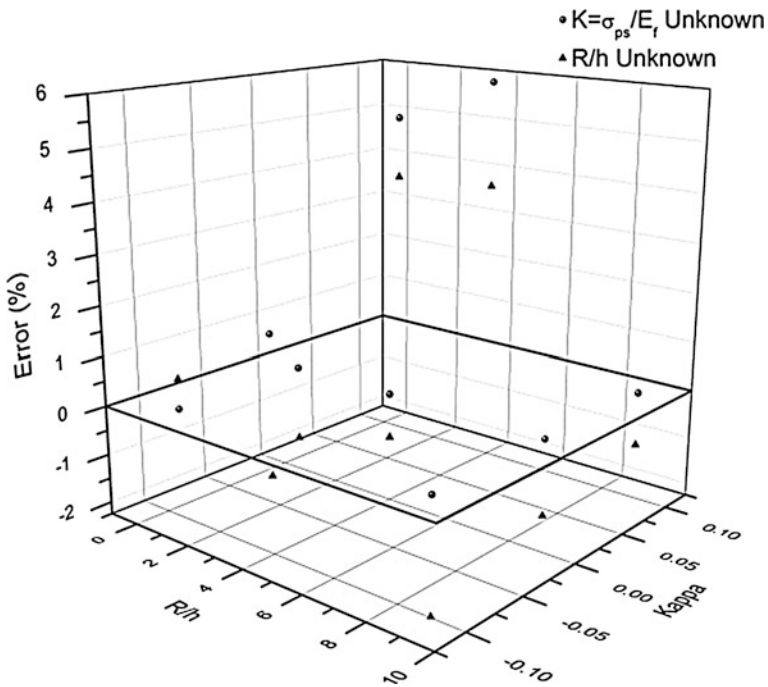


Fig. 8 Reverse analysis – error plot, $0.5 \leq R/h \leq 10.0$, $-0.10 \leq K \leq 0.10$

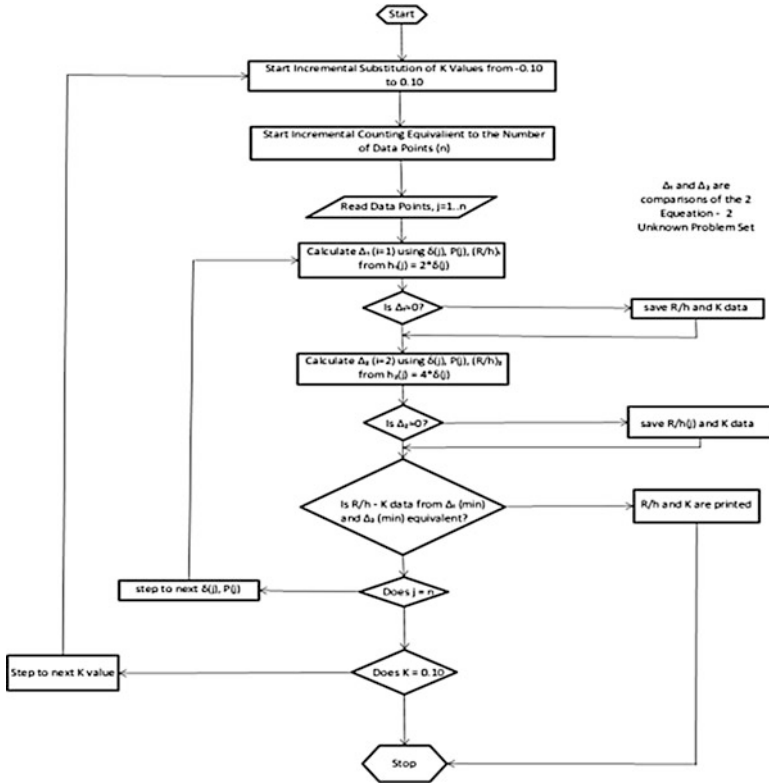


Fig. 9 General algorithm – reverse analysis

$$\frac{dP_1}{E_f \delta_1^2} - \frac{P_1 dE_f}{E_f^2 \delta_1^2} - \frac{2P_1 d\delta_1}{E_f \delta_1^3} = \frac{\partial \left[f_1 \left(\frac{E_f}{E_s}, \frac{\sigma_{ps}}{E_f} \right) \right]}{\partial \left(\frac{E_f}{E_s} \right)} d \left(\frac{E_f}{E_s} \right) + \frac{\partial \left[f_1 \left(\frac{E_f}{E_s}, \frac{\sigma_{ps}}{E_f} \right) \right]}{\partial \left(\frac{\sigma_{ps}}{E_f} \right)} d \left(\frac{\sigma_{ps}}{E_f} \right) \tag{8}$$

$$\frac{dP_2}{E_f \delta_2^2} - \frac{P_2 dE_f}{E_f^2 \delta_2^2} - \frac{2P_2 d\delta_2}{E_f \delta_2^3} = \frac{\partial \left[f_2 \left(\frac{E_f}{E_s}, \frac{\sigma_{ps}}{E_f} \right) \right]}{\partial \left(\frac{E_f}{E_s} \right)} d \left(\frac{E_f}{E_s} \right) + \frac{\partial \left[f_2 \left(\frac{E_f}{E_s}, \frac{\sigma_{ps}}{E_f} \right) \right]}{\partial \left(\frac{\sigma_{ps}}{E_f} \right)} d \left(\frac{\sigma_{ps}}{E_f} \right) \tag{9}$$

The potential errors due to indenter force and displacement ($\partial P_1/P_1$, $\partial P_2/P_2$, $\partial \delta_1/\delta_1$, $\partial \delta_2/\delta_2$) may arise from measurement errors, noise, edge effects, etc. In rearranging these terms and combining them, they can be related to the perturbation of unknowns, $\partial E_f/E_f$, $\partial E_s/E_s$, $\partial \sigma_{ps}/E_f$.

As an example, if the film modulus is known directly, then $dE_f = 0$. And Eqs. (8) and (9) can be solved directly as two independent equations with two unknowns. In effect the film prestress and substrate modulus are solved in terms of the perturbation forces and displacements:

$$\frac{d\sigma_{ps}}{E_f} = \alpha_1 \frac{dP_1}{P_1} + \alpha_2 \frac{d\delta_1}{p\delta_1} + \alpha_3 \frac{dP_2}{P_2} + \alpha_4 \frac{d\delta_2}{\delta_2} \quad (10)$$

where:

$$\alpha_1 = \frac{\Omega_1 Q_2}{w}, \alpha_2 = \frac{-2\Omega_1 Q_2}{w}, \alpha_3 = \frac{-\Omega_2 Q_1}{w}, \alpha_4 = \frac{2\Omega_2 Q_1}{w}$$

with $w = (Q_2 V_1 - Q_1 V_2)$

$$\begin{aligned} \frac{dE_s}{E_s} &= \beta_1 \frac{dP_1}{P_1} + \beta_2 \frac{d\delta_1}{\delta_1} + \beta_3 \frac{dP_2}{P_2} + \beta_4 \frac{d\delta_2}{\delta_2} \\ \beta_1 &= \frac{\Omega_1 V_2}{X}, \beta_2 = \frac{-2\Omega_1 V_2}{X}, \beta_3 = \frac{-\Omega_2 V_1}{X}, \beta_4 = \frac{\Omega_2 V_1}{X} \end{aligned}$$

with $X = \frac{E_f}{E_s} (Q_2 V_1 - Q_1 V_2) f_i \left(\frac{E_f}{E_s}, \frac{\sigma_{ps}}{E_f} \right)$ is as shown in Eq. (6). And Q_i and V_i are:

$$Q_i = \frac{\partial \left[f_i \left(\frac{E_f}{E_s}, \frac{\sigma_{ps}}{E_f} \right) \right]}{\partial \left[\frac{E_f}{E_s} \right]} = A_i C_i \left(\frac{E_s}{E_f} = B_i \right)^{C_i} + 2D_i \left(\frac{E_s}{E_f} \right) + E_i$$

$$\begin{aligned} V_i &= \frac{\partial A_i}{\partial \left(\frac{\sigma_{ps}}{E_f} \right)} \left[\frac{E_f}{E_s} - B_i \right]^{C_i} \\ &+ A_i \left\{ -C_i \left(\frac{E_f}{E_s} - B_i \right)^{C_i-1} \frac{\partial B_i}{\partial \left(\frac{\sigma_{ps}}{E_f} \right)} + \left(\frac{E_f}{E_s} - B_i \right)^{C_i-1} \ln \left(\frac{E_f}{E_s} - B_i \right) \frac{\partial C_i}{\partial \left(\frac{\sigma_{ps}}{E_f} \right)} \right\} \\ &+ \left(\frac{E_f}{E_s} \right)^2 \frac{\partial D_i}{\partial \left(\frac{\sigma_{ps}}{E_f} \right)} + \left(\frac{E_f}{E_s} \right) \frac{\partial E_i}{\partial \left(\frac{\sigma_{ps}}{E_f} \right)} + \frac{\partial F_i}{\partial \left(\frac{\sigma_{ps}}{E_f} \right)} \end{aligned}$$

with A_i, B_i, C_i, D_i, E_i $i = 1, 2$ as shown in (6) and

$$\begin{aligned} \frac{\partial A_i}{\partial \left(\frac{\sigma_{ps}}{E_f} \right)} &= a_{1_i} + 2a_{2_i} \left(\frac{\sigma_{ps}}{E_f} \right) + 3a_{3_i} \left(\frac{\sigma_{ps}}{E_f} \right)^2 + 4a_{4_i} \left(\frac{\sigma_{ps}}{E_f} \right)^3 \\ \frac{\partial B_i}{\partial \left(\frac{\sigma_{ps}}{E_f} \right)} &= b_{1_i} + 2b_{2_i} \left(\frac{\sigma_{ps}}{E_f} \right) + 3b_{3_i} \left(\frac{\sigma_{ps}}{E_f} \right)^2 + 4b_{4_i} \left(\frac{\sigma_{ps}}{E_f} \right)^3 \\ \frac{\partial C_i}{\partial \left(\frac{\sigma_{ps}}{E_f} \right)} &= c_{1_i} c_{3_i} \left(\frac{\sigma_{ps}}{E_f} - c_{2_i} \right)^{c_{s_i} - 1} \\ \frac{\partial D_i}{\partial \left(\frac{\sigma_{ps}}{E_f} \right)} &= d_{1_i} + 2d_{2_i} \left(\frac{\sigma_{ps}}{E_f} \right) + 3d_{3_i} \left(\frac{\sigma_{ps}}{E_f} \right)^2 + 4d_{4_i} \left(\frac{\sigma_{ps}}{E_f} \right)^3 \\ \frac{\partial E_i}{\partial \left(\frac{\sigma_{ps}}{E_f} \right)} &= e_{1_i} + 2e_{2_i} \left(\frac{\sigma_{ps}}{E_f} \right) + 3e_{3_i} \left(\frac{\sigma_{ps}}{E_f} \right)^2 + 4e_{4_i} \left(\frac{\sigma_{ps}}{E_f} \right)^3 \\ \frac{\partial F_i}{\partial \left(\frac{\sigma_{ps}}{E_f} \right)} &= 0 \end{aligned}$$

$i = 1, 2$ and $a_{r_i}, b_{r_i}, c_{q_i}, d_{r_i}, e_{r_i}$ ($r = 1 - 4, q = 1 - 3$) are listed in Table 1.

The terms α_j and β_j are coefficients for the perturbed errors of indentation load and displacement measurements at δ_1 and δ_2 . They can be called error sensitivity coefficients. By examining these coefficients, one may reach an understanding of the scale of error that each physical measurement could contribute to the overall error of our problem solution. As an example, if the normalized error in measuring the indentation force at δ_1 , dP_1/P_1 , is 1%, then the overall contribution to the error in the solution for the prestress is $\alpha_1\%$. And the contribution to the error in the substrate modulus is $\beta_1\%$. Therefore the smaller the error sensitivity coefficient, the less sensitive the solution algorithm is to the specific perturbation error. Moreover, the perturbation error due to measuring indentation forces at δ_1 and δ_2 would most likely have the same order of magnitude and occur together in a test. The same should be true for the errors in measuring indentation depth. Therefore error plots are shown with combined error coefficients. For example, this would then be the associated error of the prestress due to the combined perturbed normalized indentation force measurements dP_1 and dP_2 : shown as $\alpha_1 + \alpha_3$. In a similar fashion, $\alpha_2 + \alpha_4$, $\beta_1 + \beta_3$, and $\beta_2 + \beta_4$ represent the combined normalized indentation depth measurement error for prestress, combined normalized indentation force measurement error for substrate modulus, and combined normalized indentation depth measurement error for substrate modulus. Representative plots are shown in Fig. 10.

In examining these error plots, it can be seen that the error sensitivity in general is relatively low. However, there are regions where the potential error shown could be quite high. However, if two of the three variables are known, the potential areas of high error sensitivity are no longer present (this is discussed in a later part of this paper).

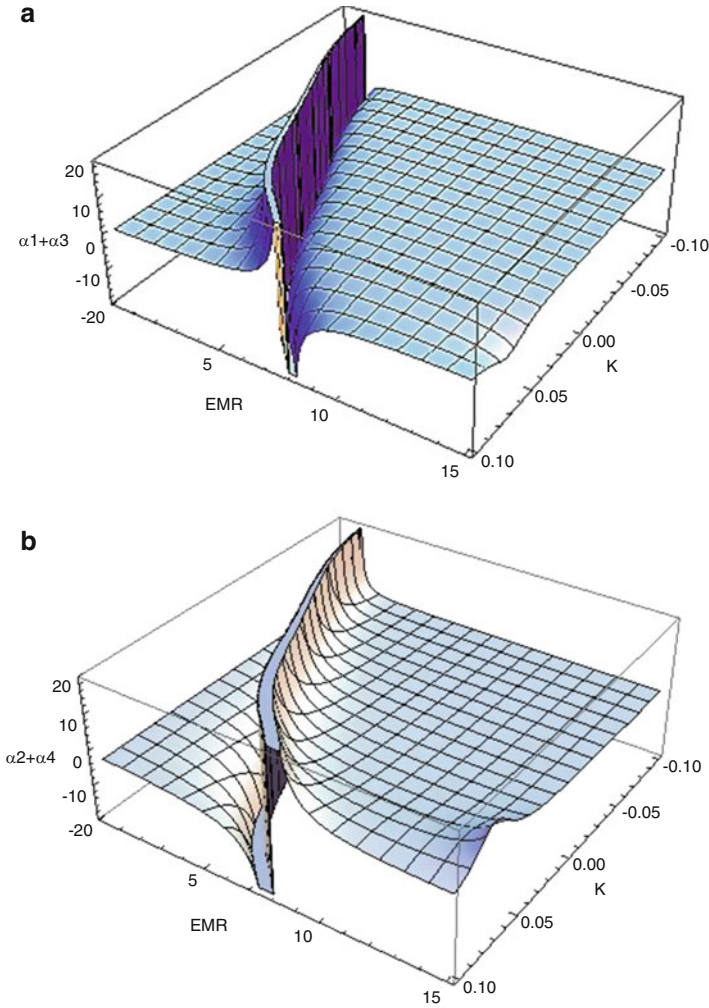


Fig. 10 Plots for (a) $\alpha_1 + \alpha_3$ and (b) $\alpha_2 + \alpha_4$ for the potential error in prestress (σ_{ps}) from measured error in indenter force, with $K = \sigma_{ps}/E_f$, $EMR = E_f/E_s$

Another example with regard to understanding measurement error sensitivity is the case where the substrate modulus is known. In this case, $dE_s = 0$. Therefore Eqs. (8) and (9) can be solved directly for the perturbed normalized film modulus and prestress:

$$\frac{dE_f}{E_f} = \xi_1 \frac{dP_1}{P_1} + \xi_2 \frac{d\delta_1}{\delta_1} + \xi_3 \frac{dP_2}{P_2} + \xi_4 \frac{d\delta_2}{\delta_2} \quad (11)$$

where:

$$\xi_1 = \frac{\Omega_1 V_2}{Y}, \quad \xi_2 = \frac{-2\Omega_1 V_2}{Y}, \quad \xi_3 = \frac{-\Omega_2 V_1}{Y}, \quad \xi_4 = \frac{2\Omega_2 V_1}{Y}$$

$$\text{with } Y = \left[\Omega_1 V_2 - Q_2 V_1 + \frac{E_f}{E_s} (V_2 Q_1 - V_1 Q_2) \right]$$

$$\frac{d\sigma_{ps}}{E_f} = \varepsilon_1 \frac{dP_1}{P_1} + \varepsilon_2 \frac{d\delta_1}{\delta_1} + \varepsilon_3 \frac{dP_2}{P_2} + \varepsilon_4 \frac{d\delta_2}{\delta_2}$$

where:

$$\varepsilon_1 = \frac{X_2}{U_1}, \quad \varepsilon_2 = \frac{-2X_2}{U_1}, \quad \varepsilon_3 = \frac{-X_1}{U_1}, \quad \varepsilon_4 = \frac{2X_1}{U_1}$$

$$\text{with } U_1 = \frac{X_2 V_1}{\Omega_1} - \frac{X_1 V_2}{\Omega_2}, \quad X_1 = \left(1 + \frac{Q_1 E_f}{\Omega_1 E_s} - \frac{V_1 \sigma_{ps}}{\Omega_1 E_f} \right),$$

$$X_2 = \left(1 + \frac{Q_2 E_f}{\Omega_2 E_s} - \frac{V_2 \sigma_{ps}}{\Omega_2 E_f} \right)$$

As a further example of the potential errors associated with the measurement of indentation data, consider the case where the prestress is known. In this case, $d\sigma_{ps} = 0$. By knowing σ_{ps} , then Eqs. (8) and (9) can be solved directly for the normalized perturbation of both the film and substrate moduli:

$$\frac{dE_f}{E_f} = \tau_1 \frac{dP_1}{P_1} + \tau_2 \frac{d\delta_1}{\delta_1} + \tau_3 \frac{dP_2}{P_2} + \tau_4 \frac{d\delta_2}{\delta_2} \quad (12)$$

where:

$$\tau_1 = \frac{Q_2 \Omega_1}{U_2}, \quad \tau_2 = \frac{-2Q_2 \Omega_1}{U_2}, \quad \tau_3 = \frac{-Q_1 \Omega_2}{U_2}, \quad \tau_4 = \frac{2Q_1 \Omega_2}{U_2}$$

$$\text{with } U_2 = Q_2 \left(\Omega_1 - V_1 \frac{\sigma_{ps}}{E_f} \right) - Q_1 \left(\Omega_2 - V_2 \frac{\sigma_{ps}}{E_f} \right)$$

and

$$\frac{dE_s}{E_s} = \eta_1 \frac{dP_1}{P_1} + \eta_2 \frac{d\delta_1}{\delta_1} + \eta_3 \frac{dP_2}{P_2} + \eta_4 \frac{d\delta_2}{\delta_2}$$

where:

$$\eta_1 = \frac{R_2}{\psi_1}, \quad \eta_2 = \frac{-2R_2}{\psi_1}, \quad \eta_3 = \frac{-R_1}{\psi_1}, \quad \eta_4 = \frac{2R_1}{\psi_1}$$

$$\text{with } \psi_1 = \frac{E_f}{E_s} \left(\frac{R_1 Q_2}{\Omega_2} - \frac{R_2 Q_1}{\Omega_1} \right), \quad R_1 = \left(1 + \frac{Q_1 E_f}{\Omega_1 E_s} - \frac{V_1 \sigma_{ps}}{\Omega_1 E_f} \right),$$

$$R_2 = \left(1 + \frac{Q_2 E_f}{\Omega_2 E_s} - \frac{V_2 \sigma_{ps}}{\Omega_2 E_f} \right)$$

Again, in solving for either normalized prestress and film modulus or normalized film and substrate modulus, the potential for errors is in general quite low. Nevertheless, there is a small risk for significant error perturbation at two extremes. The first location for potentially significant error is found where the EMR is low combined with high K . The second instance is found at a very high compressive

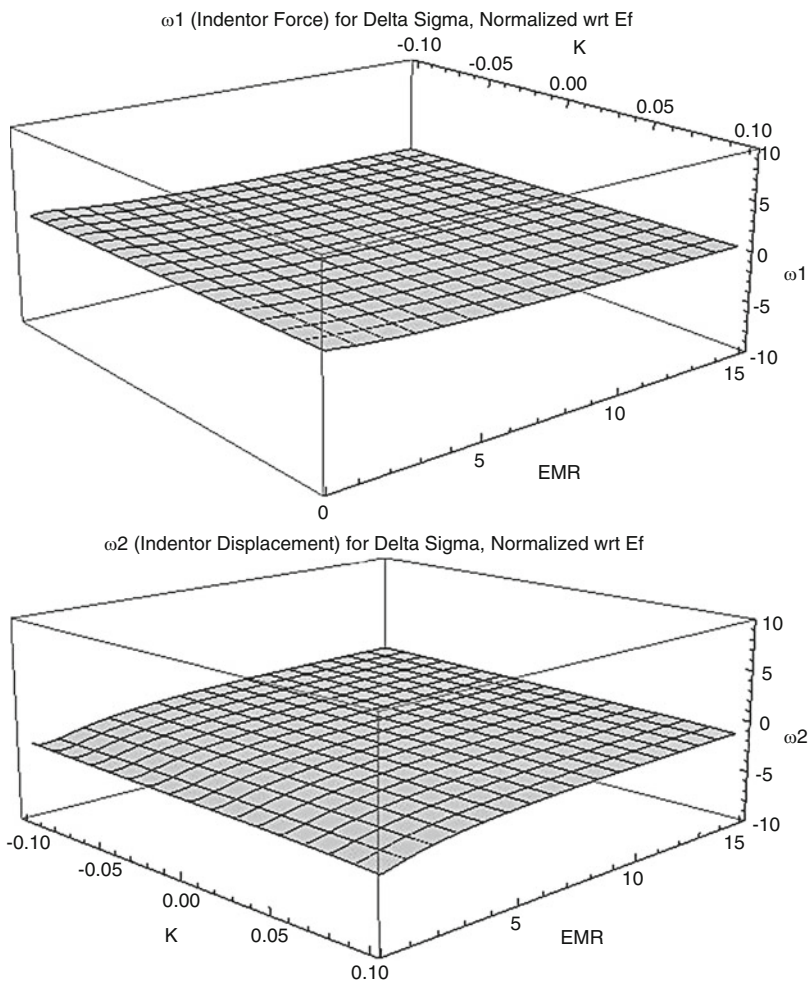


Fig. 11 Perturbation coefficient for indenter (a) force and (b) displacement measurement

prestress (low K) combined with a very high film (stiff) modulus (high EMR). In the latter case, a practical application may be that the high compressive stress may tend to cause instabilities and in doing so affect the measured indentation force and displacement measurement. In the first instance, the substrate effect is low, and the indentation behavior is dominated by bending of prestretched film (where perhaps better functional forms should be developed instead of the indentation ones).

The above analysis assumes that there are two unknowns that require solution. However there may be instances where there is only one unknown variable. As an example, if the substrate and film moduli are both known, then the potential error in prestress with respect to perturbations for force and displacement can be shown as:

$$\frac{d\sigma_{ps}}{E_f} = \omega_1 \frac{dP_1}{P_1} + \omega_2 \frac{d\delta_1}{\delta_1} \quad (13)$$

where:

$$\omega_1 = \frac{\Omega_1}{V_1}, \quad \omega_2 = -2 \frac{\Omega_1}{V_1}$$

with Ω_1 and V_1 as noted above. ω_1 and ω_2 are plotted in Fig. 11 and show that the error sensitivity coefficients are found to be stable and reasonably low, without the potential peaks noted in the two-variable error analysis.

Conclusion

Internal stress of thin films can often lead to premature failure or malfunction of a component. Meanwhile, the substrate properties are sometimes as important as that of the film. It is therefore useful to develop an effective indentation framework to probe both the substrate property and film prestress. Based on a comprehensive numerical analysis, this paper establishes algorithms to effectively deduce film/substrate elastic moduli and equibiaxial prestress. Through a systematic forward analysis, a set of fitting equations were developed for specified indentation depths at one half and one quarter of the film thickness. By incorporating spherical indentation and examining the indentation force at a moderate depth, the effects of both the film and substrate are taken into account. After the force at these specified indentation depths is measured, through the reverse analysis algorithm, one may effectively predict two of the three unknowns, film modulus, substrate modulus, or film prestress, with high accuracy. An extensive examination of error sensitivity was also made which showed that at areas of high instability (high compressive prestress), accurately measuring indentation force and depth is critical. However if two of the three properties (E_f , E_s) are known precisely, the film prestress is shown to be insensitive to minor measurement errors and should provide a robust and consistent solution. The overall strategy can be readily extended to probing elastoplastic properties in future studies.

Acknowledgments The study is supported by the National Natural Science Foundation of China (11172231), DARPA (W91CRB-11-C-0112), the World Class University Program through the National Research Foundation of Korea (R32-2008-000-20042-0), and the National Science Foundation (CMMI-0643726).

References

- J.L. Bucaille, S. Stauss, E. Felder, J. Michler, Determination of plastic properties of metals by instrumented indentation using different sharp indenters. *Acta Mater.* **51**, 1663–1678 (2003)
- X. Chen, J.J. Vlassak, Numerical study on the measurement of thin film mechanical properties by means of nanoindentation. *J. Mater. Res.* **16**, 2974–2982 (2001)
- X. Chen, Y. Xiang, J.J. Vlassak, A novel technique to measure mechanical properties of porous materials by nanoindentation. *J. Mater. Res.* **21**(3), 715–724 (2006)
- X. Chen, N. Ogasawara, M. Zhao, N. Chiba, On the uniqueness of measuring Elastoplastic properties from indentation: the indistinguishable mystical materials. *J. Mech. Phys. Solids* **55**, 1618–1660 (2007)
- Y.T. Cheng, C.M. Cheng, Analysis of indentation loading curves obtained using conical indenters. *Philos. Mag. Lett.* **77**(1), 39–47 (1998)
- Y.T. Cheng, C.M. Cheng, Scaling, dimensional analysis, and indentation measurements. *Mat. Sci. Eng.* **R44**, 91–149 (2004)
- Dassault Systèmes (SIMULIA), *ABAQUS 6.8 Documentation* (Dassault Systèmes (Simulia) Corporation, Providence, 2008)
- F. Doerner, W.D. Nix, A method for interpreting the data from depth-sensing indentation instruments. *J. Mat. Res.* **1**(4), 601–609 (1986)
- A.G. Evans, J.W. Hutchinson, On the mechanics of delamination and spalling in compressed films. *Int. J. Solids Struct.* **20**(5), 455–466 (1984)
- H. Gao, C. Cheng-Hsin, L. Jin, Elastic contact versus indentation modeling of multi-layered materials. *Int. J. Solids Struct.* **29**, 2471–2492 (1992)
- A. Gouldstone, N. Chollacoop, M. Dao, J. Li, A.M. Minor, Y.L. Shen, Indentation across size scales and disciplines: recent developments in experimentation and modeling. *Acta Mater.* **55**, 4015–4039 (2007)
- X.Z. Hu, B.R. Lawn, A simple indentation stress-strain relation for contacts with spheres on bilayer structures. *Thin Solid Films* **322**(1–2), 225–232 (1998)
- M.S. Hu, M.D. Thouless, A.G. Evans, The decohesion of thin films from brittle substrates. *Acta Metall.* **36**(5), 1301–1307 (1988)
- J.A. Mills, X. Chen, Spherical indentation on an elastic coating/substrate system: determining substrate modulus. *J. Eng. Mech.* **135**(10), 1189–1197 (2009)
- G.M. Pharr, Measurement of mechanical properties by ultra-low load indentation. *Mat. Sci. Eng. A* **253**, 151–159 (1998)
- R. Saha, W.D. Nix, Effects of the substrate on the determination of thin film mechanical properties by nanoindentation. *Acta Mater.* **50**, 23–38 (2002)
- R. Saha, Z. Xue, Y. Huang, W.D. Nix, Indentation of a soft metal on a hard substrate: Strain gradient hardening effects. *J. Mech. Phys. Solids* **49**, 1997–2014 (2001)
- S. Suresh, A.E. Giannakopoulos, A new method for estimating residual stresses by instrumented sharp indentation. *Acta Metall.* **46**(16), 5755–5767 (1998)
- J.G. Swadener, E.P. George, G.M. Pharr, The correlation of the indentation size effect measured with indenters of various shapes. *J. Mech. Phys. Solids* **50**, 681–694 (2002)
- T.Y. Tsui, J.J. Vlassak, W.D. Nix, Indentation plastic displacement field: part I: the case of soft films on hard substrates. *J. Mat. Res.* **14**, 2196–2203 (1999)
- M. Zhao, X. Chen, N. Ogasawara, A.C. Razvan, N. Chiba, D. Lee, Y.X. Gan, A new sharp indentation method of measuring the elastic-plastic properties of soft and compliant materials by using the substrate effect. *J. Mat. Res.* **21**, 3134–3151 (2006a)

-
- M. Zhao, N. Ogasawara, N. Chiba, X. Chen, A new approach to measure the elastic-plastic properties of bulk materials using spherical indentation. *Acta Mater.* **54**, 23–32 (2006b)
- M. Zhao, X. Chen, Y. Xiang, J.J. Vlassak, D. Lee, N. Ogasawara, N. Chiba, Y.X. Gan, Measuring elastoplastic properties of thin films on an elastic substrate using sharp indentation. *Acta Mater.* **55**, 6260–6274 (2007)



Experimentation and Modeling of Mechanical Integrity and Instability at Metal/Ceramic Interfaces

5

Wen Jin Meng and Shuai Shao

Contents

Introduction	154
Microscopic Mechanistic Understanding of the Failure of MCIRs	159
The Computationally Guided and Experimentally Validated ICME Framework	161
The Multiscale AIDD Simulator and VPSC-CIDD Model	163
Strategy for Improving the Mechanical Integrity of Metal-Ceramic Interfaces	165
Structure and Mechanical Instability of Metal/Ceramic Interfaces – Initial Results	167
Synthesis of Ceramic-Coating/Adhesion-Interlayer/Substrate Systems and Structural/Mechanical Characterization of Coating/Substrate Interfacial Regions	167
Mechanistic Understanding on the Shear Instability of Metal/Ceramic Interface Provided by Simulations	190
Summary	204
References	205

Abstract

Controlling the mechanical integrity of metal/ceramic interfaces is important for a wide range of technological applications. Achievement of such control requires a number of key elements, including establishing appropriate experimental protocols for quantifying mechanical response of metal/ceramic interfacial regions under well-defined loading conditions, understanding how interfacial compositional and structural characteristics impact such interfacial mechanical response, and elucidating unit interface physics and predicting interfacial mechanical response via development of multiscale physics-based models. Achieving this combined testing, understanding, and modeling will ultimately lead to effective

W. J. Meng (✉) · S. Shao

Department of Mechanical and Industrial Engineering, Louisiana State University, Baton Rouge, LA, USA

e-mail: wmeng1@lsu.edu; sshao@lsu.edu

© Springer Nature Switzerland AG 2019

G. Z. Voyiadjis (ed.), *Handbook of Nonlocal Continuum Mechanics for Materials and Structures*, https://doi.org/10.1007/978-3-319-58729-5_50

153

control of mechanical integrity of metal/ceramic interfaces and true interfacial engineering through targeted modification of the interfacial composition and structure.

Major breakthroughs in the improvement of interfacial mechanical integrity can be enabled by understanding and controlling key physical factors, including interfacial architectural and chemical features governing the mechanical response of metal/ceramic interfacial regions (MCIRs), thus leading to unprecedented interfacial mechanical performance that meets/exceeds the demands of future applications. Guided by a multiscale integrated computational materials engineering (ICME) framework, the mechanical integrity of MCIRs can be substantially improved by a variety of architectural and chemical enhancements/refinements. Recent research efforts by the authors aim to provide a fundamental, physics-based understanding of the failure mechanisms of MCIRs by constructing a novel, multiscale, computation-guided, and experiment-validated ICME framework. Interfacial refinements to be explored within this framework include addition of alloying impurities as well as geometrical features such as multilayered and stepped interfacial architectures. The findings can then be consolidated into a high fidelity, experiment-validated, micro- and mesoscale modeling tool to significantly accelerate the discovery-design-implementation cycle of advanced MCIRs. In this chapter, we summarize some preliminary results on shear failure and instability of various metal/ceramic interfacial regions, outline the theoretical background of this research thrust, and identify challenges and opportunities in this area.

Keywords

Metal/ceramic interfaces · Microscale mechanical testing · Integrated computational materials engineering (ICME)

Introduction

Understanding and controlling mechanical integrity of solid/solid interfaces have remained a scientific challenge over the last three decades. The continued interest in this subject stems from the positive and significant impacts to wide ranging technological applications such understanding and control can bring. Notable technological examples include thermal barrier coatings (TBCs) for turbine components (Darolia 2013), hard ceramic coatings for machining tools (Holmberg and Matthews 2009), and hard/lubricious coatings for mechanical components such as gears and bearings (Mercer et al. 2003; Kotzalas and Doll 2010). In TBC systems with a ZrO₂-based ceramic top coating layer, various solid/solid interfaces are present: a metal/metal interface as represented by the bond coat/superalloy substrate interface; a metal/ceramic interface as represented by the thermally grown oxide (TGO)/bond coat interface; and a ceramic/ceramic interface as represented by the TBC top-layer/TGO interface. Failures at these interfaces constitute important mechanisms limiting the durability of TBC systems (Evans et al. 2008).

In coatings for machining tools and mechanical components, the application of thin hard ceramic coatings (CHCs) on metallic substrates likewise requires sufficient mechanical integrity of the interfacial regions between thin ceramic layers and metallic substrates. In the application of thin CHCs onto machining tools and mechanical components, adequate mechanical integrity of the metal/ceramic interfacial region (MCIR) constitutes a “go/no-go” requirement dictating whether a particular application can or cannot be adopted. In the past three decades, CHCs with thickness less than 10 μm have been implemented in many surface engineering applications. In such applications, thin interlayers are often used to promote adhesion between the ceramic coating layers and the substrates, forming thin-ceramic-layer/adhesion-interlayer/substrate sandwich structures (Jiang et al. 2003). A major failure mode for CHC/substrate systems, often with catastrophic consequences, is the mechanical failure of the MCIR, see for example Fig. 1a. Due to the present lack of understanding regarding key physical factors governing mechanical response of MCIRs, “good” or “bad” interfaces can only be distinguished through testing under actual application conditions, a comparative example is shown in Fig. 1a, b. The materials paradigm of synthesis-structure-property-performance circle is thus largely missing at the present time when it comes to MCIRs.

Manipulating interfacial chemistry and architecture may improve its shear/tensile failure stress. For instance, the introduction of low-energy ledges/steps as well as point defects has been demonstrated to control both the shear and dislocation-nucleation resistance of a sharp bi-metal interface (Zhang et al. 2016; Skirlo and Demkowicz 2013). Multiphase nano-adhesion interlayers can perform better than a compositionally sharp MCIR, as shown in Fig. 1a, b (Jiang et al. 2003). Intriguingly, strengthening effects are also observed in metal/ceramic nanolaminates, where the interfaces are arranged in a parallel fashion with nanoscale spacing (Li et al. 2014; Bhattacharyya et al. 2011). Recent work has shown that when the individual layer thickness of a metal/ceramic nano-laminate is very small, e.g., 2 nm, the ceramic layers co-deform plastically with the metal layers (Wang et al. 2017). In this case, under compressive loading normal to the interfaces, the nanolaminated metal/ceramic composites can reach an ultimate compressive stress of 4.7 GPa and a total compressive strain of 13% before the onset of instability, exhibited through the formation of shear bands (Fig. 1c–e). When the individual layer thickness is large, however, plastic co-deformation cannot be achieved which results in cracking in the ceramic layers. While such anecdotal evidence shows promise for engineering the mechanical integrity of MCIRs through interfacial architectural and chemical design, systematic research in this direction has not been carried out and engineering realization has seldomly been demonstrated. Little quantitative data exist at present on the characterization of interfacial mechanical response of coating/metal-adhesion-layer/substrate sandwich structures and on tailoring the chemistry and architecture of the MCIR to influence the interfacial mechanical response. Engineering of a particular coating/interlayer/substrate system has therefore largely proceeded in a trial-and-error manner, necessitating testing under actual application conditions, which is both time consuming and expensive.

Currently, understanding the mechanical response of interfacial regions of thin-ceramic-layer/metal-interlayer/substrate systems is hampered by a lack of

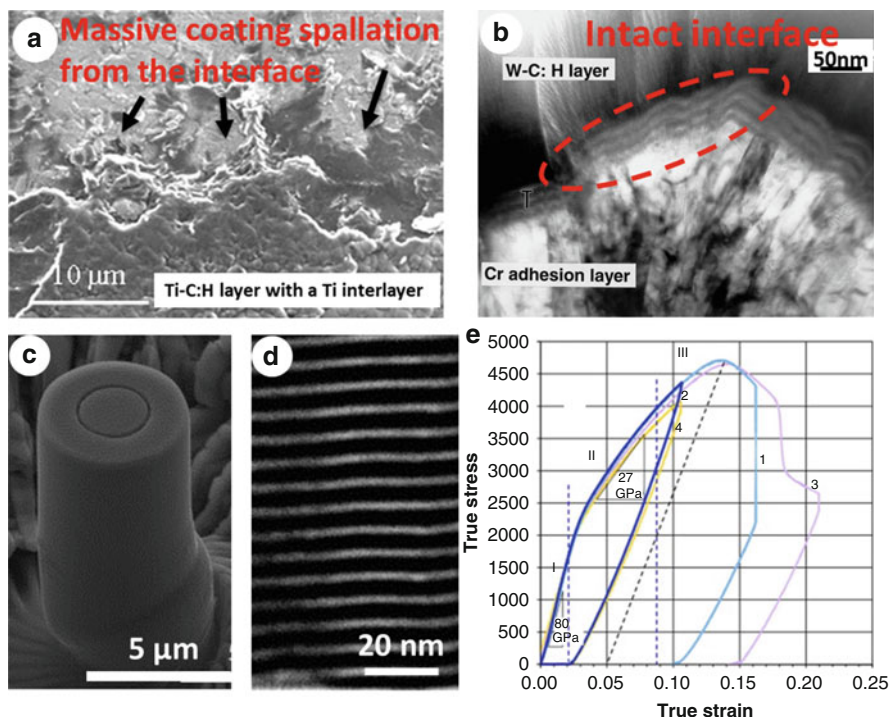


Fig. 1 (a) Contact fatigue testing of Ti-containing amorphous hydrogenated carbon (Ti-C:H) coated gears resulting in massive spallation failure at the Ti-C:H/Ti-interlayer interface; (b) intact metal/ceramic interfacial region (MCIR) after contact fatigue testing of W-containing amorphous hydrogenated carbon (W-C:H) coated gears. Here, the MCIR has a more complex architecture, with a Cr adhesion layer in contact with the steel substrate, followed by nano-laminated interlayers in contact with the W-C:H top coating layer (Jiang et al. 2003); (c–e) exceptional mechanical properties of metal/ceramic nano-laminates (Bhattacharyya et al. 2011)

experimental measurement protocols that are capable of generating quantitative data on interfacial mechanical response that can be easily interpreted. This understanding is further limited by a lack of physics-based modeling and simulations on the mechanical response of coating/substrate interfaces. Particularly lacking is an integration of experiments and modeling/simulation efforts, where information obtained from experiments probing the mechanical response of interfacial regions under sufficiently simple loading conditions can be coupled with outputs from multiscale modeling/simulation efforts, so as to develop deeper understandings of the key mechanisms governing mechanical response of interfacial regions under different loading conditions.

Established testing methods for adhesion of thin layers on substrates have mostly been developed to work at the macroscale. For example, adhesion tests under tensile loading conditions (e.g., see ASTM International (2002, 2012, 2014) and shear loading conditions (e.g., see ASTM International (2005, 2007, 2013) both

require attachment of the loading device to the thin coating layer by application of a thin epoxy layer. The strength of epoxies is typically insufficient to fail the strong interfacial regions between thin ceramic coatings and metallic substrates typically found in coatings for machining tools and mechanical components, thus resulting in the inapplicability of such macroscale tensile pull-off and shear adhesion tests to engineered thin ceramic coatings. The use of epoxy also limits the range of temperatures at which testing is conducted (Brown 1994, 1995).

While the scratch testing protocol (e.g., ASTM International (2015)) has been widely applied to evaluating adhesion of hard coatings on metallic substrates, the conversion of the so-called critical load data to quantitative measures of coating/substrate interfacial mechanical response has been recognized to be extremely difficult if not impossible. This is due to the many different and complex physical processes that can occur in a scratch test, including energy dissipation within the ceramic coating layer due to coating cracking and energy dissipation within the coating layer or the substrate due to plastic deformation, as well as the complex stress field surrounding the moving scratch indenter (Bull 2001). The critical load measured from scratch testing does not only depend on the adhesion between the coating and the substrate, as it should be in the ideal case (Kutilek and Miksovsky 2011). The complications listed above are responsible for reports in the literature regarding the dependence of the measured critical load on a wide range of parameters of the entire coating/substrate system, including the coating's thickness, hardness, and internal stress state, the hardness of the substrate, and the loading conditions (Lukaszkoewicz et al. 2011).

At the present time, the laser spallation test offers, arguably, the best experimental methodology for obtaining a quantitative measure of the interfacial tensile strength between thin films and substrates (Gupta et al. 1992, 2003). In the laser spallation test, an energetic laser beam is incident upon the substrate side of a thin-film/substrate specimen and sets up a compressive stress pulse in the substrate, the reflection of which from the free surface of the film generates tension at the substrate/film interface, and causes film spallation from the substrate when this interfacial tensile stress reaches a critical magnitude. While laser spallation tests have yielded quantitative measures of interfacial tensile strength (Gupta et al. 1994), this testing methodology is associated with limitations due to specimen geometry (typically flat-wafer type specimens) and substrate absorption. Past laser spallation measurements on ceramic-layer/metal-adhesion-layer/metal-substrate systems have also shown an apparent dependence of the interfacial strength on the thickness of metal interlayer (Gupta et al. 1994). Observation of such a dependence of strength value on interlayer thickness indicates that the laser spallation testing protocol also may not yield a value correspondent with an ideal interfacial strength, and suggests the need for an improved understanding of the exact nature of interfacial failure induced by the laser-induced tensile interfacial stress pulse, even though the laser spallation measurement is considered well established and well studied.

The small thickness of the coating layer in ceramic-coating/metal-interlayer/substrate systems, ranging typically from $\sim 0.2 \mu\text{m}$ to $\sim 10 \mu\text{m}$, makes their mechanical coupling to macroscale external loading systems difficult. This difficulty

is in part responsible for the lack of progress with respect to the development of reliable and quantitative mechanical testing protocols for the mechanical response of coating/substrate interfacial regions. In this regard, the development of focused ion beam (FIB) micro-/nano-scale machining capabilities (Volkert and Minor 2007) and in situ instrumented micro-/nano-scale mechanical testing capabilities over the last decade opens up new avenues for obtaining quantitative data as well as more detailed observations related to mechanical failure of coating/substrate interfacial regions. Preliminary microscale mechanical testing results presented in this chapter offer some examples for this research direction.

Early simulations related to the mechanical response of solid/solid interfaces, such as *ab initio* density functional theory (DFT) calculations on metal/ceramic interfaces performed two decades ago, were limited to ideal interfacial configurations without taking into account potential defect configurations at/near interfaces (Smith et al. 1994). The dramatic improvements in computing power occurred over the last two decades have enabled multiscale simulations related to solid/solid interfaces, from DFT to molecular dynamics (MD), to dislocation dynamics (DD), and to grain-level plasticity models, such as crystal-plasticity finite element models (CPFEM) or visco-plastic self-consistent (VPSC) models. Hierarchical multiscale modeling schemes that integrates interfacial deformation physics across a broad spectrum of lengths scales have been developed (Zbib and Diaz de la Rubia 2002; Groh et al. 2009), which have the capabilities of addressing interfacial mechanics problems that are much more realistic and relevant to experimental conditions. We believe that recent advances in microscale mechanical testing capabilities and development of multiscale physics-based modeling/simulation tools offer an opportunity to deepen the current understanding of the mechanical response of solid/solid interfaces. In the particular case of ceramic-coating/substrate interfacial regions, this confluence of micro-/nano-mechanical testing and multiscale modeling/simulations can offer new insights into the critical elements governing the mechanical integrity of interfacial regions between thin ceramic layers and substrates. Preliminary simulation results presented in this chapter again offer some examples.

In this chapter, we first briefly summarize the current understanding in the existing literature regarding the mechanisms of mechanical instabilities in metal/ceramic interfaces (section “[Microscopic Mechanistic Understanding of the Failure of MCIRs](#)”). In section “[The Computationally-Guided and Experimentally-Validated ICME Framework](#),” we propose a novel integrated computational materials engineering (ICME) framework, currently being implemented by the authors, aimed to improve mechanical integrity of the metal/ceramic interfaces. Details regarding its approach, concerning both computation and experimentation, are discussed. In section “[Structure Mechanical Instability of Metal/Ceramic Interfaces – Initial Results](#),” we illustrate a specific effort under such a combined experimentation-simulation framework – the evaluation of shear and compression instabilities of CuN/Cu/Si and CrN/Ti/Si interfacial regions. In this effort, instrumented compression testing was performed, with concurrent scanning electron microscopy (SEM) observations, on microscale cylindrical pillars fabricated from vapor deposited

CrN/Cu/Si(001) and CrN/Ti/Si(001) specimens. Such technique allowed a much more detailed observation of how the interfacial regions fail under load. Concomitant simulations were carried out on Ti/TiN interfaces, employing MD in combination with DFT. The simulation results shed light on the influence of the Ti/TiN interface on the mechanical behavior of the Ti atomic layers nearby, in both shear and tensile strength. The combined DFT/MD simulations illustrate the link between length scales (from the atomic- to nano-) and were effective in probing structural features' influence on mechanical response of realistic semi-coherent interfaces. It is worth noting that this chapter is not meant to be a thorough literature review. Instead, the intention of this chapter is to present a novel ICME-based approach, with a specific example, to evaluate and/or improve the mechanical stability of the MCIRs.

Microscopic Mechanistic Understanding of the Failure of MCIRs

The close proximity of reactive nonmetallic element strongly impacts the electronic structure inside the metal layer near the chemical interface, resulting in weaker theoretical shear and tensile strengths (Zhang et al. 2017; Yadav et al. 2015). Recently, by combining experimental observations made during in situ microscale mechanical testing with MD simulations and DFT calculations, the present authors have provided new insights regarding mechanical failures of thin-ceramic-layer/metal-interlayer/substrate systems (Zhang et al. 2017). DFT calculations performed on the Ti/TiN interface suggest a weakening effect of the metal/ceramic chemical interface to its adjacent metal atomic monolayers in both shear and tension, parallel and normal to the interface. This is evidenced by a reduced generalized stacking fault energy (GSFE) profile as well as work of adhesion (WoA) on the (0001) slip planes in Ti near the chemical interfaces. Similar observation has also been observed for Al/TiN and Al/VN interfaces (Yadav et al. 2015). As a result of the weakened energetic characteristics, the misfit dislocation network (MDN) formed in such planes exhibits increasingly planar dislocation cores (Shao et al. 2013, 2014, 2015), higher dislocation mobility (Hirth and Lothe 1982), as well as vanishing pinning effects of the nodes (dislocation intersections) (Zhang et al. 2017; Shao et al. 2015). On the other hand, MD simulations suggest that the free energy of the MCIR is also at a minimum when the MDN locates away from the chemical interface, coinciding with the MDN position where the shear strength is a minimum (Zhang et al. 2017). This explains the experimental observation that the eventual shear failure occurs not at, but adjacent to, the chemical interface (Zhang et al. 2017).

When the MCIR is subjected to a normal stress, deformation behavior and failure mechanisms are vastly different depending on the length scale of the microstructures. For isolated individual interfaces within the MCIRs (e.g., direct coating without interlayers or laminated interlayers with large layer thickness, e.g., $d > \sim 500$ nm), gliding lattice dislocations in metal interact with individual sharp interfaces and form dislocation pile-up (Wang et al. 2017; Hirth and Lothe 1982; Wang and Misra 2011). The significant tensile stress at the tip of a dislocation-pile

up may lead to formation of cracks (Meyers and Chawla 2007) for tensile loading normal to the MCIR. Note that due to the reduced theoretical tensile strength, as measured by the WoA, in the metal near the chemical interface in the MCIR, the tensile cracks may preferentially form in the metal layers leading to delamination (Fig. 2a). Under compressive loading, the local normal-to-interface tensile stress component at the tip of the pile up counteracts the applied compressive stress and may be less severe for opening cracks. In such a case, the parallel-to-interface tensile stress component at the tip of the pile-up superimposed with the tensile stress caused by the elasto-plastic deformation incompatibility between ceramic and metal layers may induce cracks perpendicular to the interface (Fig. 2b). When multiple chemical interfaces are adjacent to each other, such as in the MCIRs within nano-laminated interlayers, either of the following situations may apply: (1) the smaller interfacial spacing ($d < 50$ nm) inhibits the formation of dislocation pile-up, and the propagation of threading dislocation within the metal layer (or confined layer slip (CLS)) is the governing mechanism; (2) the interfacial spacing is so small ($d < 10$ nm) that CLS is inhibited and nucleation/emission of dislocations from the MDN is preferred. If the thickness of the ceramic is small, e.g., in the range of only several nanometers, the strong dislocation dipoles form on both sides of the ceramic layer may activate slip without cracking (Fig. 2c) (Li et al. 2014; Bhattacharyya et al. 2011; Wang and Misra 2014). For slightly larger ceramic layer thicknesses,

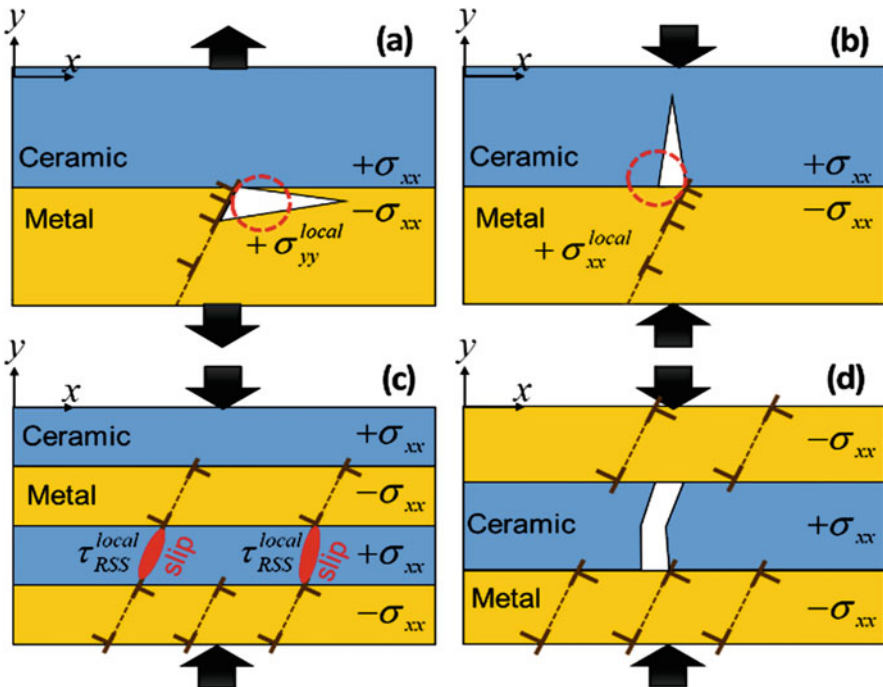


Fig. 2 Schematic illustration of the deformation and failure mechanisms experienced by MCIRs

cracks normal to interfaces may still form, due to the tensile stress in ceramic layer originated from the elasto-plastic deformation incompatibility between ceramic and metal layers (Fig. 2d).

The Computationally Guided and Experimentally Validated ICME Framework

The macroscopic failure of the CHC/substrate system is often governed by the atomic scale defect evolution events at the MCIR (Li et al. 2014; Bhattacharyya et al. 2011; Zhang et al. 2017; Wang and Misra 2014; Yang et al. 2017). For instance, in response to compressive/tensile loading normal to the MCIR, the MDN may nucleate and emit lattice dislocations, which will in turn interact with gliding lattice dislocations. The pile-up of dislocations at an interface may lead to the opening of cracks and hence failure of the CHC. Under shear loading, the MDN may nucleate interface dislocations that propagate parallel to the MCIR. Alternatively, as recently demonstrated by the present authors (Zhang et al. 2017), the MDN may glide collectively in response to an applied shear stress. Both deformation mechanisms lead to shear failure of the MCIR. Therefore, variations in the architecture and/or chemistry at interfaces may hugely impact the defect activities (Wang and Misra 2011; Shao and Wang 2016; Wang et al. 2014; Salehinia et al. 2014, 2015) and consequently improve the mechanical response of MCIRs.

Given the multiscale nature of mechanical failure in MCIRs, we envision a multiscale computationally guided and experimentally validated ICME framework (Fig. 3). At the atomic scale, using density functional theory (DFT) and molecular dynamics (MD) simulations, tensile and shear strength of candidate MCIR with various chemical (metal/ceramic composition as well as interfacial dopants) and architectural (interfacial steps and ledges (Zhang et al. 2016; Hirth and Pond 1996; Henager et al. 2004; Hirth et al. 2006) and arrangement of individual metal/ceramic interfaces in an MCIR) modifications are explored to identify potentially strong interfaces. Selected MCIR designs can then be experimentally implemented through state-of-the-art plasma-assisted vapor phase synthesis, mechanically tested at the microscale under different loading conditions, and characterized through microscopy and spectroscopy techniques. Experimental results, in turn, can validate the computational MCIR design. The so-obtained deformation physics is passed to and implemented in the microscale atomistically informed interface dislocation dynamics (AIDD) model. The output of the validated AIDD model, including the dislocation evolution law and the constitutive relations, is incorporated into a mesoscale visco-plastic self-consistent continuum interface dislocation dynamics (VPSC-CIDD) model (Lyu et al. 2015, 2016, 2017). The output of the VPSC-CIDD model, i.e., the texture evolution and stress-strain response are then validated against experimentation. This VPSC-CIDD model is expected to predict the mechanical/failure strength of MCIRs with arbitrary geometries, such as within metal matrix composites (MMCs), and not limited to the

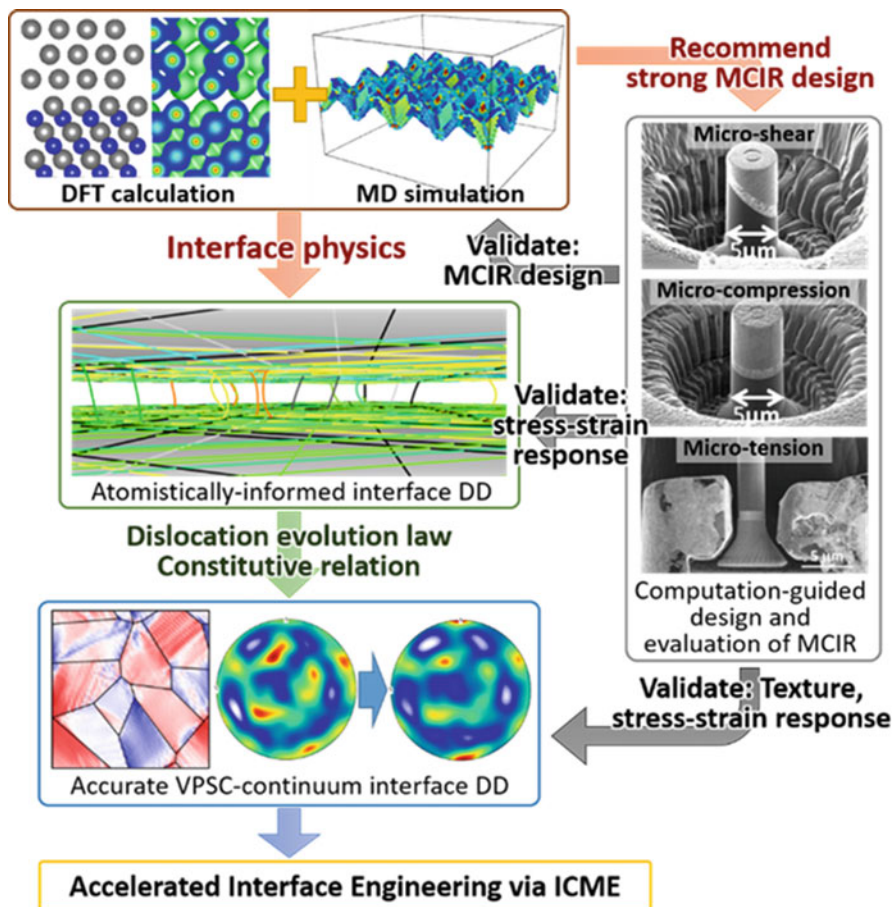


Fig. 3 Illustration of the multiscale, computationally guided, and experimentally validated ICME framework

conventional layered structures (Lyu et al. 2015, 2016, 2017; Lebensohn and Tomé 1993, 1994). Such an ICME framework is expected to advance basic understanding of the physics governing mechanical response of metal/ceramic interfaces, and will have general applicability including, but not limited to, design of CHC systems.

Through regulation of interfacial chemistry and architecture, the structure-property relationship of mechanically strong MCIRs can be explored, established, and validated by combining multiscale modeling/computation with vapor phase synthesis, microscale mechanical testing, and multiscale materials characterization. AIDD and VPSC models can then be advanced to incorporate the structure and deformation physics of the MCIR, which in turn, will enable accelerated design and implementation of strong MCIRs. The multiscale computationally guided and experimentally validated ICME framework will empower a significantly accelerated

and more cost-effective design-implementation cycle (encompassing the discovery, validation, modeling, and design) for advanced MCIRs, leading to next generation applications.

The Multiscale AIDD Simulator and VPSC-CIDD Model

The current AIDD model incorporates interface deformation mechanisms and interface structural features in classical dislocation dynamics (DD) models (Zbib and Diaz de la Rubia 2002; Wang et al. 2014; Akasheh et al. 2007; Han and Ghoniem 2005; El-Awady et al. 2008; Ghoniem et al. 2000, 2002; Ghoniem and Han 2005; Zbib et al. 1998). Two distinct types of dislocations are modeled – interface dislocation and lattice dislocation. When lattice dislocations approach within a critical distance to an interface, they are captured by the interface and converted to interface dislocations; their Burgers vectors are conserved. If an interface dislocation with near-parallel line sense is present, the incoming lattice dislocation may react with it and form a new interface dislocation, if the reaction leads to a lower dislocation line energy. On the other hand, when a “weak” interface with relatively low shear strength is considered, in response to the stress field of an incoming lattice dislocation, interface dislocation loops may nucleate, which result in localized shear. The interface is modeled as a special, piecewise flat plane, allowing the glide and climb of interface dislocations, depending on the Burgers vector, i.e., when $\mathbf{b}_{inter} \cdot \hat{\mathbf{n}}_{inter} = 0$, the dislocation is glissile. Otherwise, it can only move by climb. The motion of interface dislocations is governed by a set of equations of motion very similar to lattice dislocations. However, due to the unique characteristics and structure of the interfaces, the parameters for the motion of the interface dislocations, e.g., drag coefficients and vacancy/interstitial migration energies, are different from those for the lattice dislocations and are therefore necessary to be calibrated using atomistic simulations and experimental observations. The reaction between two critically close interface dislocations is also permitted if both reactant dislocations have near-parallel line sense and if the resulting dislocation has lower line energy.

As was discussed previously, the plastic deformation in the nano-laminated interlayers is mainly carried out through the CLS mechanism as well as the nucleation and subsequent propagation of lattice dislocations from the interfaces. MD simulations (Fig. 4a) (Beyerlein et al. 2013a, b; Zhang et al. 2013; Shao et al. 2017) have revealed that a nucleation event is favored if the nucleation source (interface line defect) is parallel to the trace of the slip plane to be activated (Situation 1 in Fig. 4b) (Shao et al. 2017). If no existing interface dislocation is aligned with the trace of slip plane, local reorientation (Situation 2 in Fig. 4b) of line defects and local reaction (Situation 3 in Fig. 4b) between line defects may occur to produce segments of defects aligned with the traces of slip planes (Shao et al. 2017). Accordingly, this interface physics has been implemented in the AIDD model. Further, the nucleation of the lattice dislocation is explicitly treated as a thermally activated process using a Monte Carlo type of approach (Fig. 4c) (Shao et al. 2017).

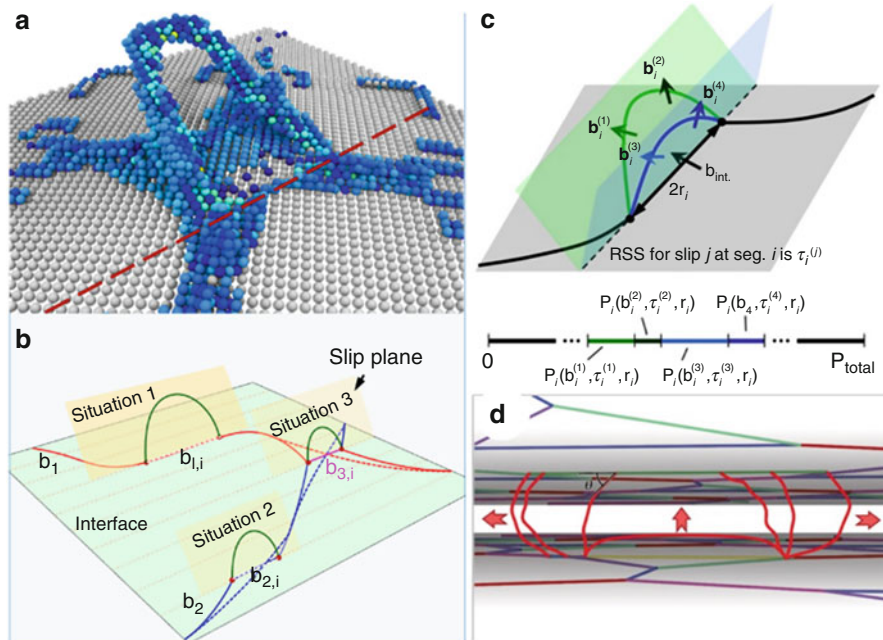


Fig. 4 Nucleation and subsequent propagation of dislocations as modeled in AIDD (Shao et al. 2017)

The newly nucleated dislocation loop bows out, driven by the externally applied stress superimposed by the local stress field of the dislocations (Fig. 4d) (Shao et al. 2017). The nucleated dislocation propagates within the layer. Once it touches the adjacent interface, it is deposited on the interface. If the angle between a dislocation segment and the interface is small, such that $\theta < \theta_c$, it will be continuously deposited on the interface, and the remaining dislocation will “thread” within the layer. Using this nucleation scheme, slip transmission can be captured. Differing from nucleation at an intrinsic interface dislocation, slip transmission occurs through nucleation from an extrinsic interface dislocation. In such an event, the Burgers vector of the residual interface dislocation is significantly reduced, leading to a significantly higher probability.

The AIDD model is limited in both time (micro seconds) and length (below 1 μm) scales and is therefore unable to predict the deformation behavior of materials in the meso- and macroscales. Its advantage is to predict the evolution law of dislocations, including the evolution of dislocation density (both mobile and immobile) as well as rates of dislocations’ nucleation and cross-slip events, to name a few. Such rules are relied upon by the higher length scale models, such as crystal plasticity finite element methods (CP-FEM) (Groh et al. 2009; Marin 2006; Casals and Forest 2009; Miller et al. 2004) as well as the VPSC model (Lebensohn and Tomé 1993, 1994; Wang et al. 2010). The VPSC-CIDD model is based upon the

conventional VPSC model with expanded capabilities in the description of plasticity (Lyu et al. 2015, 2016, 2017). Ignoring elastic deformation, the VPSC-CDD model determines the critical resolved shear stress on slip system α based on a Bailey-Hirsch type of relation (Lyu et al. 2015, 2016, 2017): $\tau_H^\alpha = Cb\mu \sum_m \Omega^{\alpha m} \sqrt{\rho_{TS}^m}$, where C is a numerical constant, b and μ are the Burgers vector and shear modulus of the crystal, $\Omega^{\alpha m}$ is the dislocation interaction matrix between current slip system α and the system m , ρ_{TS}^m is the density of the statically stored dislocations (SSD) on slip system m . ρ_{TS}^m is described based on several different contributions, such as dislocation multiplication, pinning and de-pinning of dislocations, annihilation, cross-slip. The characteristics of all of these processes can be explicitly obtained by AIDD simulations (Lyu et al. 2015, 2016, 2017). The outputs of the VPSC-CIDD model (Fig. 3) include the mesoscale stress-strain relation, the local stress and strain distribution, as well as the texture evolution – all of which can be verified from mechanical testing and material characterization experiments, e.g., electron backscatter diffraction (EBSD) and transmission electron microscopy (TEM).

Strategy for Improving the Mechanical Integrity of Metal-Ceramic Interfaces

Based on the existing literature (Yadav et al. 2015; Lin et al. 2017; Sun et al. 2017) and the recent work by the present authors (Zhang et al. 2017), it is clear that there exists a weak plane in the vicinity of the chemical interface for many MCIRs containing sharp interfaces. Such weak interaction planes give rise to weak shear as well as tensile strengths, which may lead to premature interfacial failure, e.g., failure of interfacial region in CHC/metal substrate systems. On the other hand, the utilizations of the nanolaminate adhesion interlayers, although substantially improves the strength and ductility of the MCIR, may still suffer from interface cracking if the individual layer thickness is not carefully controlled. Therefore, strategies for improving the strength and ductility of MCIRs have to address the failure mechanisms identified above. Three potential paths of structural and chemical modifications to MCIRs are being considered, requiring the understanding of key physical factors controlling the mechanical response in each case, and consequently developing the best approaches and solutions to engineer the mechanical integrity of MCIRs (Fig. 5):

1. **Adding interfacial impurities/solute-atoms** to improve the theoretical shear and tensile strength of a MCIR (Fig. 5a). In doing so, the characteristics of the MDN, i.e., the core structure, mobility, and pinning effects of the dislocations/nodes, may be altered. Such an alteration may result in either a stronger shear resistance of MDN at its minimum energy location or a change in the minimum energy location that corresponds to a higher shear strength;
2. **Utilizing a nanolaminate type adhesion interlayer** that composes of alternating metal and ceramic layers with thickness of only a few nanometers (Fig. 5b) to improve the compressive strength (stress before the onset of instability, such as

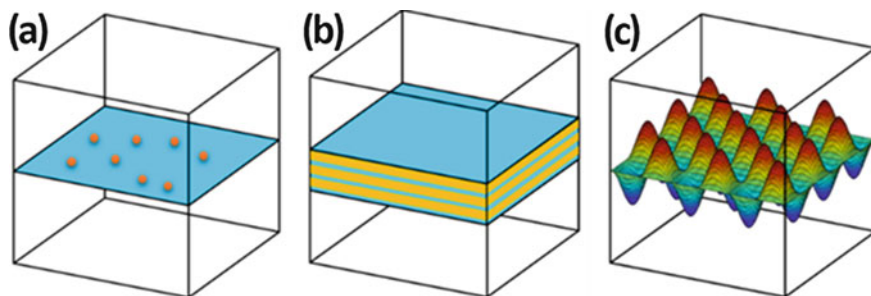


Fig. 5 Three potential paths for improving the strength and ductility of metal/ceramic interface regions. (a) Interface with impurity, (b) nanolaminate interface, (c) 3D structured interface

shear bands or cracks) and ductility of the MCIR. The thickness ratio between the metal and ceramic layers need to be investigated for how it impacts the overall stress response.

3. **Designing an interface with pronounced micro- and/or nanoscale roughness** to improve the shear strength of the MCIR (Fig. 5c). The roughness can be achieved by introducing interface steps/islands with low energy ledges, which will locally transform the parallel-to-interface shear stress into normal-to-interface tensile/compressive stress.

Note that the three potential paths may be combined with each other to achieve an optimum performance.

The extent of possible interfacial strength improvement is expected to vary depending on choice of material/impurity as well as the length scale of the interfacial architectural features. Indeed, implementing such modifications to MCIRs through a random experimental search will be inefficient, if not downright impossible. The presently described multiscale, computationally guided, and experimentally validated ICME framework is expected to generate mechanistic understanding, through multiscale simulation, of how the three paths affect the deformation/failure behavior of MCIRs. This will include (1) using DFT to explore a combination of metal/ceramics that yields the strongest pristine MCIR, (2) using DFT to probe the effect of the various solute atoms on the GSFE and WoA for various metal/ceramic interfaces, (3) parameterizing modified embedded atom method (MEAM) interatomic potentials, and (4) using existing and parameterized MEAM potentials to perform MD studies to investigate the microscopic deformation and failure mechanisms in MCIRs associated to each modification path taken.

The knowledge/insight gained through modeling/simulation can then guide the design of experiments in devising novel interface chemistries/structures. The experimental valuation of the mechanical response of MCIRs in conjunction with detailed multiscale materials characterization will provide feedback to the modeling/simulation efforts. In the meantime, the interface physics obtained through MD simulations can be used to advance the existing AIDD model and VPSC-CIDD

model. A closed feedback loop is therefore formed: the in situ and ex situ experimental results are used to validate the VPSC-CIDD model, which in turn will be used to discover, design, and implement novel interfaces.

Structure and Mechanical Instability of Metal/Ceramic Interfaces – Initial Results

In this section, we provide one instance of effort under such a combined experimentation-modeling/simulation framework – the investigation of shear instability of the interfacial regions of two ceramic-coating/metal-adhesion-layer/substrate systems: CrN/Cu/Si and CrN/Ti/Si. The failure of the interfacial regions was measured quantitatively and observed directly through instrumented compression of cylindrical micropillars with concurrent scanning electron microscopy observations. The results indicated, for the first time to our knowledge, that shear failure of the interfacial region occurred in two stages: an initial shear deformation of the entire metal interlayer followed by an unstable dynamic shear-off close to one metal/ceramic interface. The experimentally observed dynamic shear-off was suggested to be concomitant with the metal/ceramic interface going from being “locked,” with no relative displacement between materials on the two sides of the interface, to being “unlocked,” with significant relative displacements. DFT and MD studies on a related metal/ceramic interface, Ti/TiN, provided further insights into this behavior. It was shown, again for the first time to our knowledge, that a weak interaction plane exists in the metal layer near the chemical interface in a coherent Ti/TiN structure. Consequently, the free energy as well as the theoretical shear strength of the semi-coherent Ti/TiN interface was found to depend on the physical location of the misfit dislocation network (MDN). The minimum energy and strength of the interface occur when the MDN was near, but not at the chemical interface.

Synthesis of Ceramic-Coating/Adhesion-Interlayer/Substrate Systems and Structural/Mechanical Characterization of Coating/Substrate Interfacial Regions

Vapor Phase Deposition of Coating/Substrate Systems

Deposition of ceramic coatings onto the surfaces of machining tools and mechanical components needs to satisfy several general requirements. For surface engineering purposes, thin CHCs are often used, e.g., transition metal carbides and nitrides, with typical physical characteristics of high melting temperatures, high elastic stiffness, high hardness, and high enthalpy of formation (Toth 1971). For satisfactory mechanical and tribological performance, the thin CHC layers deposited onto the substrates need to be fully dense and well adherent to the substrate. Such ceramic layers are typically deposited onto the substrate surface by vapor phase deposition, usually through either chemical vapor deposition (CVD), physical

vapor deposition (PVD), or some form of hybrid vapor deposition technique (Vossen and Kern 1991). Deposition of refractory ceramic layers by conventional CVD methods usually involves high substrate temperatures, often exceeding the temperature limitations of the metallic substrates and causing degradation of the substrate mechanical properties. To deposit fully dense refractory ceramic layers at low temperatures typically requires the input of additional energies to the growth surface to counterbalance the lack of adatom surface mobility during growth and is typically achieved through some form of ion beam or plasma assist (Itoh 1989).

Thin transition metal nitride layers are used as examples of CHCs, the deposition of which was carried out at close to room temperature in an ultra-high-vacuum hybrid PVD/CVD deposition tool. This deposition tool housed a 13.56 MHz inductively coupled plasma (ICP) and multiple balanced magnetron sputter sources (Meng et al. 1999). The sputter sources faced the center of the deposition chamber, with a base pressure of $<3 \times 10^{-9}$ Torr. Cleaned silicon and metal substrates were first placed into a load lock, evacuated to $\leq 3 \times 10^{-7}$ Torr, and then transferred to a holder placed at the center of the deposition chamber. The substrates were rotated at ~ 12 rpm during deposition. Pure metal (99.95%+) targets were operated in the dc current-controlled mode. The entire deposition sequence occurred in ~ 10 mTorr of Ar (99.999%+). The substrates were first subjected to an Ar ICP etch for ~ 5 min at a bias voltage of -50 V. Immediately after etching, elemental metal adhesion interlayers were deposited onto the substrate by sputtering the elemental metal targets in Ar with ICP assist. Deposition of metal nitride coating layers, ~ 5 μm in thickness, occurred immediately after the metal adhesion interlayer deposition in an Ar/N₂ (99.999%+) ICP. To ensure a reasonable deposition rate at close to stoichiometry, the input N₂ flow was kept close to but below the pressure hysteresis point. During metal nitride deposition, a total input ICP power of 1000 W was applied and an electrical bias voltage of -30 V to -100 V was applied to the substrate. Additional details on the methodology of using low pressure high density plasma assisted vapor phase deposition to form fully dense CHC layers have been presented elsewhere (Meng et al. 1999, 2000; Meng and Curtis 1997).

Structural, Compositional, and Mechanical Characterization of Vapor Deposited Thin Film/Coating Specimens

Structures of metal-nitride/metal-adhesion-interlayer/substrate specimens were characterized by combining X-ray diffraction (XRD), scanning electron microscopy (SEM), Ga⁺ focused ion beam (FIB) sectioning, and TEM. A PANalytical Empyrean system with Cu K α radiation was used for XRD measurements. Diffraction patterns were obtained with specimens mounted on a χ - ϕ - ψ - θ stage. Diffraction patterns in the glancing incidence geometry, the symmetric θ - 2θ geometry, and the ω -rocking curve geometry are obtained with an incident beam graphite mirror and a Pixel3D detector. X-ray pole figure data were obtained with an incident beam double-cross slit and a scintillation detector.

Scanning imaging with electron- or ion-induced secondary electrons (SE/ISE) and Ga⁺ FIB milling were carried out on an FEI Quanta3D Dual-Beam FEG instrument, which housed a 30 kV field-emission electron source, a 30 kV

high-current Ga⁺ ion source, an OmniProbe for specimen lift-out, and an EDAX X-ray energy-dispersive spectroscopy (EDS) attachment. A JEOL JEM2011 microscope operated at 200 kV was used for TEM examinations. Cross-sectional TEM specimens were made by Ga⁺ ion sectioning into rectangular sections, lift-out using an in situ OmniProbe, gluing by Ga⁺ catalyzed Pt deposition onto a TEM grid, followed by Ga⁺ ion thinning at 30 kV and 5 kV. Final specimen thinning/cleaning was performed on a Gatan PIPS II ion polishing system using an Ar⁺ ion beam at 500 eV or less. Further details regarding Ga⁺ ion beam sectioning and TEM specimen lift-out have been described elsewhere (Chen et al. 2014a). Additional TEM specimens were made without the final Ar⁺ ion thinning/cleaning step. With respect to characterizing the grain structure, no significant difference was found between specimens with or without the final low energy Ar⁺ thinning/cleaning step.

Compositional characterization was carried out by X-ray photoelectron spectroscopy (XPS) on a Kratos AXIS165 spectrometer with monochromatic Al K α excitation. Composition quantification was obtained from raw XPS spectra using factory supplied sensitivity factors. Prior to XPS spectra collection, the specimen surface was sputter cleaned with an Ar⁺ ion beam for ~20 min, with the Ar⁺ ion beam set at 4 kV and 15 mA.

Instrumented indentation was carried out at room temperature on a Nanoindenter XP system, using a three-sided pyramidal Berkovich diamond indenter. The indenter's Young's modulus and Poisson's ratio are 1170 GPa and 0.07, respectively. The machine compliance and the projected indenter tip area as a function of the indenter contact depth was calibrated using a factory supplied fused silica standard following the Oliver-Pharr method (Oliver and Pharr 1992). The Young's modulus and Poisson's ratio for the fused silica standard were taken, respectively, to be 72 GPa and 0.18, independent of the indenter contact depth. The calibration covered a contact depth range from 40 nm to 2100 nm. Raw indentation loading and unloading curves were obtained in the load-controlled mode using a constant loading and unloading time of 15 s, with a 30 s load hold at the maximum load, L_{\max} . Multiple load versus indenter displacement, L - d , curves were obtained at one d_{\max} value, and d_{\max} was varied to obtain a complete set of indentation data.

In what follows, structural, compositional, and mechanical characterization of polycrystalline CrN and Cu thin films deposited onto Si(001) substrates is presented in some detail as an example. The same characterization methodology was applied to all metal-nitride/metal-adhesion-layer/substrate systems. In particular, XRD studies of CrN/Ti/Si specimens showed that the Ti interlayer is hcp in structure and has a strong texture with the Ti hexagonal basal plane parallel to the Si substrate.

A series of single-layer Cu thin films were deposited onto Si(001) substrates. Cu deposition occurred with an Ar ICP assist at a total ICP input power of 1000 W. A substrate bias of -50 V was applied during deposition. The Cu film thickness was controlled through the deposition time. Another series of bilayer CrN/Cu thin films were deposited onto Si(001) substrates. Cu interlayers were first deposited onto Si(001) in pure Ar with ICP assist, followed immediately by deposition of CrN top layers in an Ar/N₂ mixture, also with ICP assist at a total ICP input power

of 1000 W. During CrN deposition, a substrate bias of -100 V was applied. To ensure CrN deposition at close to stoichiometry, the input N_2 flow was kept above the pressure hysteresis point.

Typical results of morphological and compositional characterizations from CrN/Cu/Si(001) specimens are illustrated in Fig. 6. Figure 6a shows an SE image of a FIB cross section perpendicular to the original specimen surface, created by Ga^+ ion milling. A Pt protection layer was laid on top of the specimen surface prior to ion milling by Ga^+ catalyzed deposition from an organometallic Pt precursor. Figure 6a shows in sequence the Pt protection layer, the CrN top layer, the Cu interlayer, and the Si substrate. The thicknesses of the CrN and Cu layers were measured, respectively, to be $\sim 5.5 \mu m$ and ~ 810 nm, with typical scatter of $\pm 5\%$ based on repeat measurements. Figure 6b shows a plan-view SE image of the as-deposited CrN top surface, with surface roughness typical of vapor deposited polycrystalline

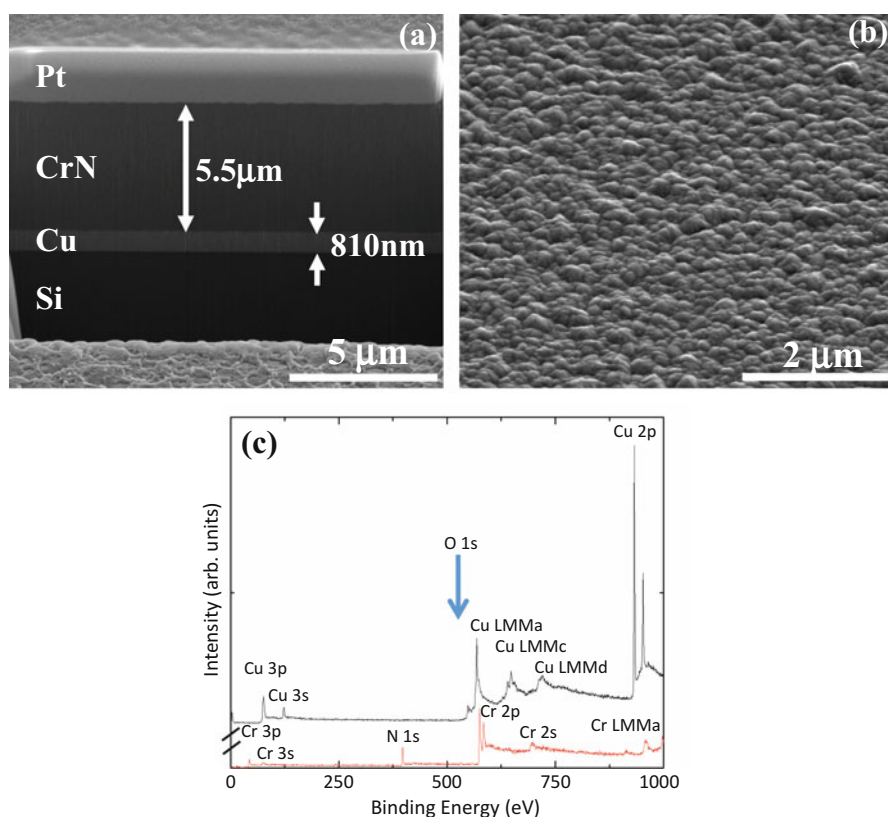


Fig. 6 Morphological and compositional characterization of a typical CrN/Cu/Si(001) specimen: (a) a FIB cross sectional SE image, (b) a 52° tilted plan-view SE image of the CrN top layer, (c) XPS survey spectra collected from top layers of a typical Cu/Si specimen (top trace) and a typical CrN/Cu/Si specimen (bottom trace). In both cases, the oxygen and carbon impurity levels were at or below the XPS detection limit of ~ 1 at.%

columnar refractory ceramics. Both the cross section and the plan-view images indicate that the CrN top layer is fully dense. Figure 6c shows XPS survey spectra collected from top layers of a typical Cu/Si specimen and a typical CrN/Cu/Si specimen. The Cu spectrum shows peaks consistent with Cu3p, Cu3s, and Cu2p emissions. Signal at the O1s and C1s positions, at binding energies of 532 eV and 284 eV, is not above the noise level. The CrN spectrum shows peaks consistent with Cr3p, Cr3s, Cr2p, Cr2s, and N1 s emissions. Signal at the O1s and C1s positions is also not above the noise level. The Cr:N ratio is determined to be 51:49. XPS data thus show that the Cu layers consist of elemental Cu and that the CrN layers consist of Cr and N at close to the stoichiometric ratio. The O and C impurity levels in both Cu and CrN layers are below the XPS detection limit of ~1 at.%. Similar compositional characterization data showed that all metal-nitride layers deposited were close to stoichiometry and all elemental metal layers deposited were free of O and C contamination above the XPS detection limit.

Results of XRD characterization of the structure of CrN layers in CrN/Cu/Si(001) specimens are illustrated by data shown in Fig. 7. The glancing incidence XRD pattern, shown in Fig. 7a, collected from one typical CrN/Cu/Si(001) specimen at an X-ray beam incidence angle $\omega = 2.5^\circ$, shows diffraction peaks only from the CrN top layer. All diffraction peaks can be indexed to a cubic structure with lattice parameter $a = 4.18 \text{ \AA}$, close to the bulk B1-CrN lattice parameter of 4.15 \AA (Toth 1971). The θ - 2θ XRD pattern, shown in Fig. 7b and collected from the same specimen, shows the Si(004) diffraction peak from the Si substrate and one minor Cu(220) diffraction peak from the Cu interlayer in addition to all diffraction peaks indexed to B1-CrN, with $a = 4.18 \text{ \AA}$. Taken together, glancing incidence and θ - 2θ XRD data are consistent with the CrN film having the B1-NaCl structure, being polycrystalline, and with preferential alignment of CrN<111> and CrN<200> along Si[001], the growth direction.

The level of residual stress within the CrN top layer was estimated from analysis of the glancing incidence XRD data shown in Fig. 7a. When the X-ray incidence angle was fixed at $\omega = 2.5^\circ$ and the detector was scanned to obtain diffraction signals in the range of $20^\circ < 2\theta < 120^\circ$, the angle between the scattering vector \vec{k} and the specimen surface normal \vec{n} , Ψ , is given by $\Psi = \theta - \omega$ (Perry et al. 1996). The CrN lattice parameter, a_{CrN} , was calculated from the multitude of B1-CrN diffraction peaks shown in Fig. 7a and plotted versus $\sin^2 \Psi$ in Fig. 7c. Assuming an equal biaxial residual stress σ_R existing within the CrN layer, the variation of a_{CrN} as the Ψ angle varies relates in the following way to σ_R (Chen et al. 2014a; Cullity and Stock 2001):

$$a_\Psi = \left[\frac{1 + \nu}{E} \sigma_R \right] a_o \sin^2 \Psi + \left[1 - \frac{2\nu}{E} \sigma_R \right] a_o. \quad (1)$$

In Eq. 1, E and ν are, respectively, the Young's modulus and Poisson's ratio of the CrN layer, and a_o is the bulk lattice parameter of CrN. A linear least squares fit to the data shown in Fig. 7c yielded a slope of $-0.022 \pm 0.017 \text{ \AA}$ and an

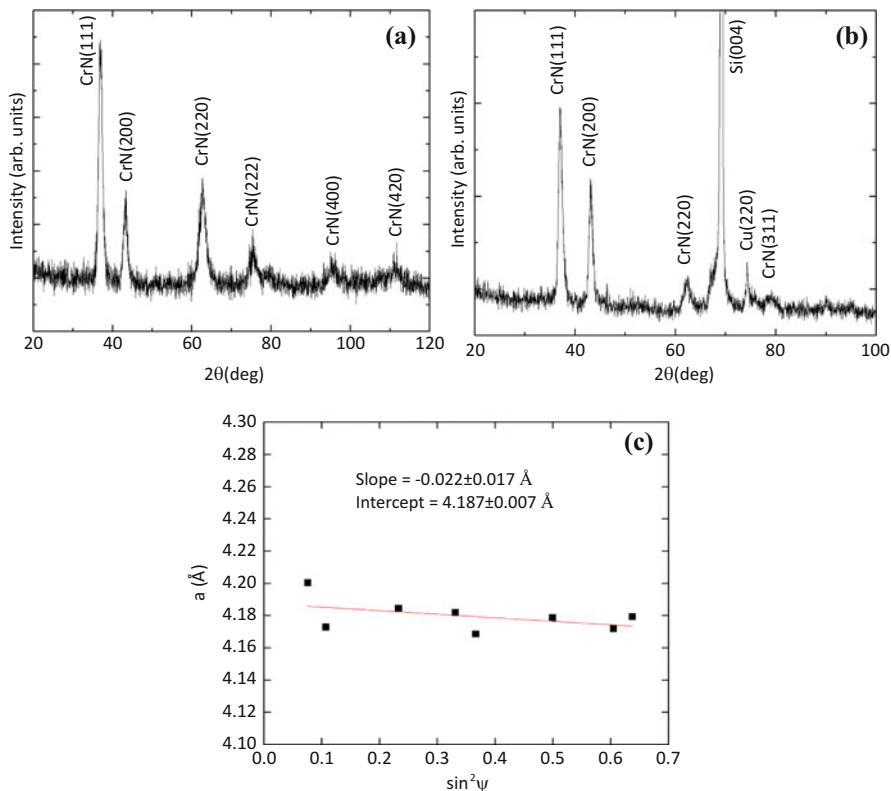


Fig. 7 XRD examination of a typical CrN/Cu/Si(001) specimen: (a) glancing incidence ($\omega = 2.5^\circ$) diffraction pattern, (b) θ - 2θ diffraction pattern, (c) $\text{Sin}^2\Psi$ analysis of the glancing incidence data shown in (a)

intercept of $4.187 \pm 0.007 \text{ \AA}$. The negative slope of the $\text{sin}^2\Psi$ plot indicates that the residual stress within CrN is compressive. Independent of the values of E and ν , the fact that the slope uncertainty is comparable to the slope itself indicates that the level of residual stress is moderate. Values of E for CrN reported in the literature vary widely from $\sim 200 \text{ GPa}$ to above 400 GPa (Kral et al. 1998; Chen et al. 2004). Taking a_0 to be 4.15 \AA , ν to be 0.2 (Sue et al. 1994), and E to be 400 GPa (Holleck 1986), the fitted slope of the $\text{sin}^2\Psi$ plot yields $\sigma_R = -1.8 \pm 1.4 \text{ GPa}$. It is noted that since $E = 400 \text{ GPa}$ is a high value for CrN, this σ_R value represents an upper bound estimate (e.g., taking $E = 300 \text{ GPa}$ would result in a σ_R value of $-1.4 \pm 1.0 \text{ GPa}$). Both the magnitude and the large uncertainty associated with the σ_R value again indicate that the level of compressive residual stress within the CrN layer is moderate. A moderate residual stress level within CrN makes it easier to deposit thick CrN top layers, e.g., that shown in Fig. 6a. Diffraction data similar to those shown in Fig. 7 were obtained from other CrN/Cu/Si(001) specimens.

Results of XRD characterization of the structure of Cu/Si(001) specimens are illustrated by data shown in Fig. 8. The glancing incidence XRD pattern, shown

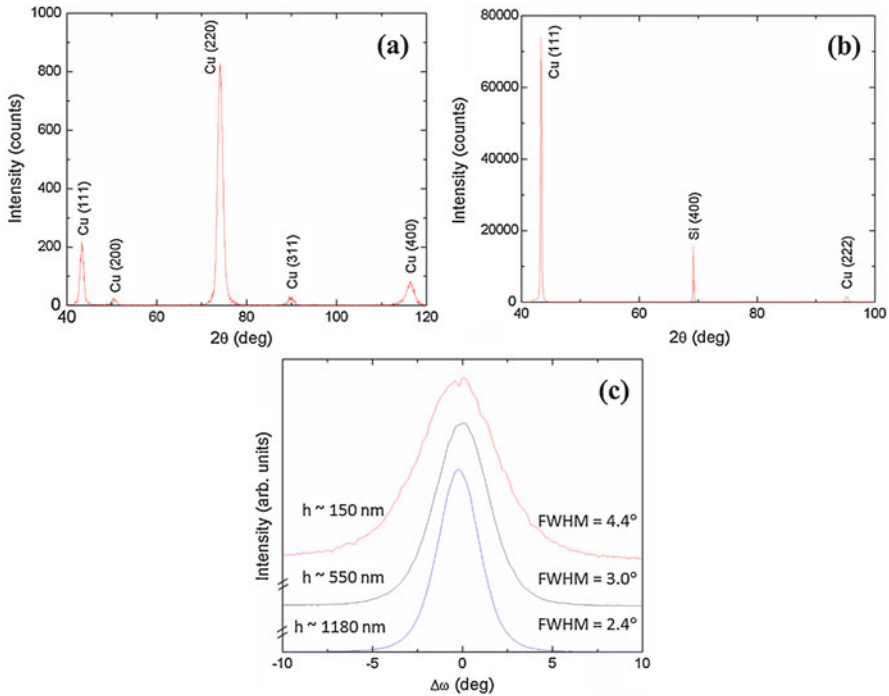


Fig. 8 XRD examination of a typical Cu/Si(001) specimen with a Cu film thickness of ~ 1180 nm: (a) glancing incidence ($\omega = 2.5^\circ$) diffraction pattern, (b) θ - 2θ diffraction pattern, (c) Cu $\langle 111 \rangle$ rocking curves from three polycrystalline Cu layers on Si(001). The Cu film thickness values are indicated on top of the rocking curves

in Fig. 8a, collected at an beam incidence angle $\omega = 2.5^\circ$ from one Cu/Si(001) specimen with Cu thickness of ~ 1180 nm, shows diffraction peaks which can all be indexed to an fcc structure with lattice parameter a of 3.617 \AA , close to the bulk fcc-Cu lattice parameter of 3.615 \AA (Straumanis and Yu 1969). The θ - 2θ XRD pattern, shown in Fig. 8b and collected from the same specimen, shows the Si(004) diffraction peak from the Si substrate and only fcc-Cu (111) and (222) diffraction peaks. The glancing incidence and θ - 2θ XRD data are consistent with the Cu film having the fcc structure, being polycrystalline, and with strong preferential alignment of Cu $\langle 111 \rangle$ along Si[001]. The Cu(111) ω -rocking curve data are shown in Fig. 8c. Measured rocking curve width decreases monotonically from 4.4° at the Cu film thickness of ~ 150 nm to 3.0° at ~ 550 nm and to 2.4° at ~ 1180 nm. The texture in Cu films was shown in more detail by background corrected Cu(111) pole figure data shown in Fig. 9, collected from three different Cu films deposited on Si(001), with total film thickness values of ~ 150 nm, ~ 550 nm, and ~ 1180 nm. In all cases, the Cu(111) diffraction intensity peaks along the growth direction and drops to $<10\%$ of the peak intensity at directions deviating from the growth direction for $>5^\circ$. Furthermore, the pole figure data clearly show in-plane rotational symmetry

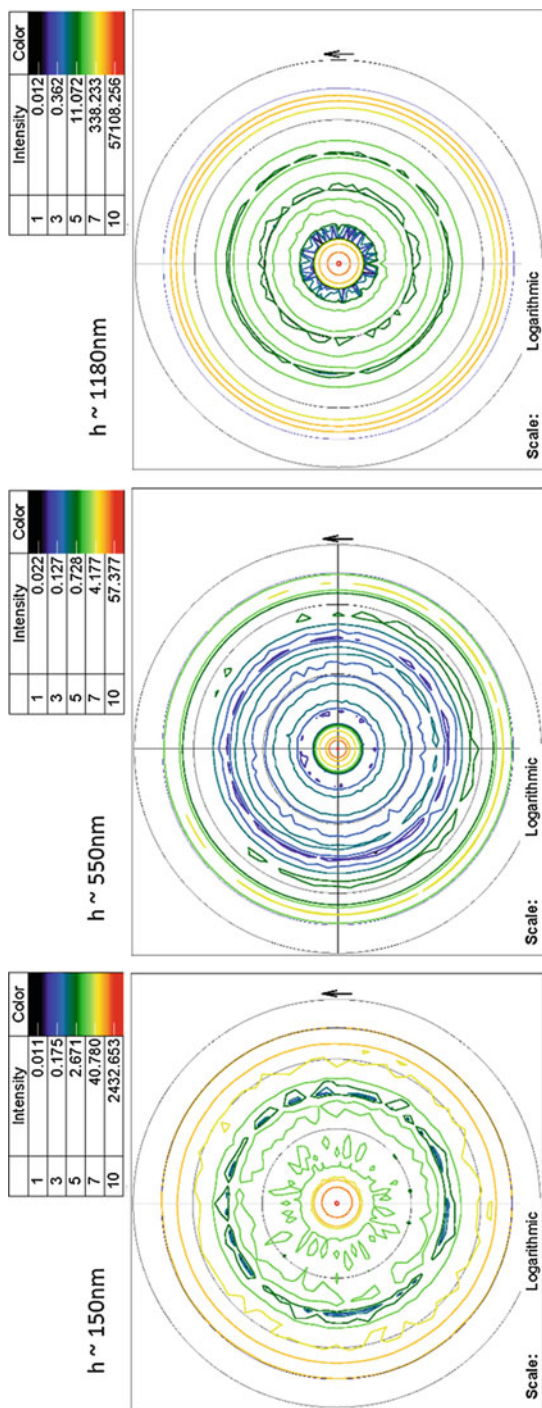


Fig. 9 Background corrected Cu <111> pole figures from three polycrystalline Cu layers deposited on Si(001). The Cu film thickness values are indicated on top of the pole figures

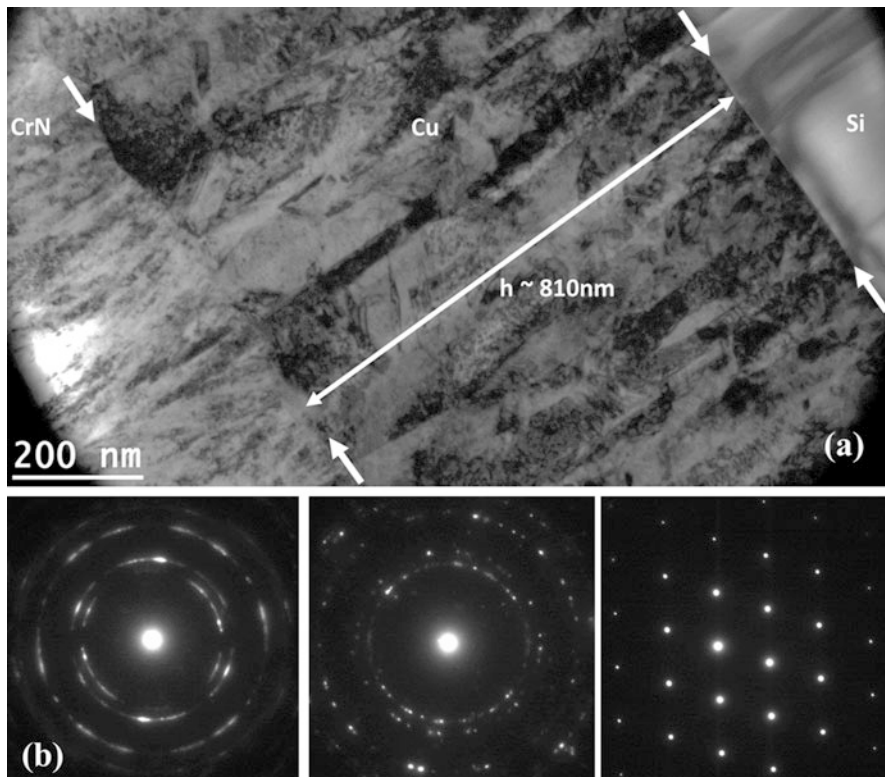


Fig. 10 TEM examination of a typical CrN/Cu/Si(001) specimen: (a) a cross sectional BF image across the entire Cu interlayer. The CrN/Cu and Cu/Si interfaces are indicated by white arrows; (b) typical SAED patterns collected from the (left to right) CrN, Cu and Si layer

of the Cu(111) diffraction intensity, at all film thicknesses probed. The pole figure data are consistent with the Cu(111) ω -rocking curve data shown in Fig. 8c. Taken together, X-ray diffraction data shown in Figs. 8 and 9 indicate that the Cu films are fcc in structure and have almost perfect fiber texture with Cu<111> along Si[001] to within a few degrees. The Cu films are polycrystalline with random in-plane orientation.

Grain structure of CrN and Cu thin films was further characterized by TEM. Figure 10 shows typical TEM characterization results of CrN/Cu/Si(100) specimens across the entire Cu interlayer. A cross-sectional TEM bright-field (BF) image, shown in Fig. 10a, was obtained from the CrN/Cu/Si(001) specimen whose morphology was shown in Fig. 6. The specimen was oriented with the Si substrate in the [110] zone axis direction. Both CrN and Cu layers are polycrystalline. Both CrN/Cu and Cu/Si interfaces appear clean, without indication of interdiffusion and reaction. For this specimen, the Cu interlayer thickness was measured from Fig. 10a to be ~ 810 nm, which was consistent with the value measured by FIB

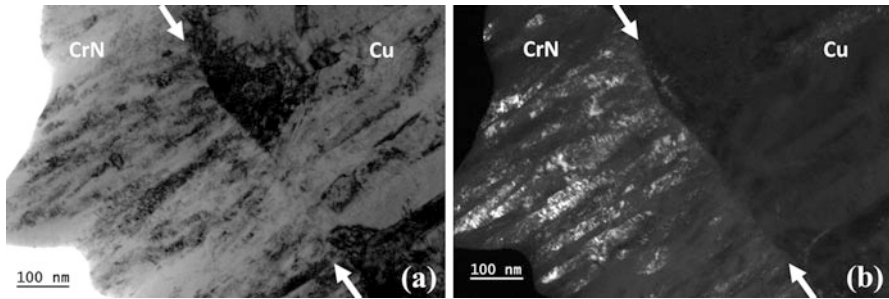


Fig. 11 Cross sectional TEM examination of the CrN top layer: (a)/(b) a BF/DF image pair of the same area. White arrows indicate the location of the CrN/Cu interface

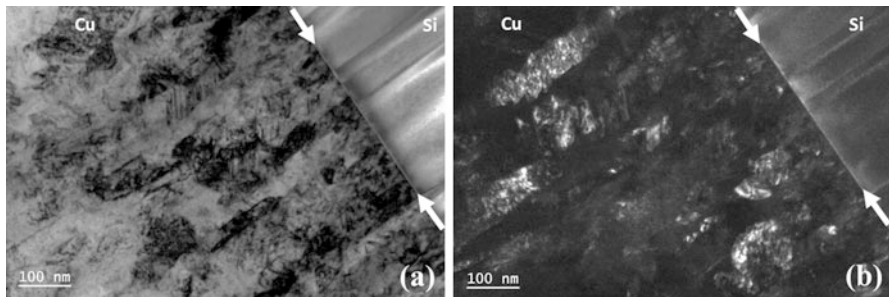


Fig. 12 Cross sectional TEM examination of the Cu interlayer: (a)/(b) a BF/DF image pair of the same area. White arrows indicate the location of the Cu/Si interface

cross sectioning in Fig. 6a. Figure 10b shows selected area diffraction patterns (SADPs) obtained, respectively, from the CrN top layer, the Cu interlayer, and the Si substrate areas. The SADP from Si shows the Si[110] zone axis diffraction pattern. The SADPs from Cu and CrN are consistent with the structure of Cu being fcc and that of CrN being B1-NaCl. Figure 11 shows a typical TEM bright-field/dark-field (BF/DF) image pair of the CrN top layer near the CrN/Cu interface. The CrN layer has a pronounced columnar structure, with columns with widths ~ 50 nm and large length/width ratios. The columnar structure of CrN shown in cross sectional TEM is consistent with the rough surface morphology observed in Fig. 6. The CrN layer appears fully dense, without indication of intercolumnar voids. Figure 12 shows a typical TEM BF/DF image pair of the Cu interlayer near the Cu/Si interface. The Cu layer consists of a random mixture of columnar and near equi-axed grains. The presence of nano-twins within the columnar Cu grains is visible in the BF image, with twin spacing < 100 nm. Similar image features were observed from Cu films at different thicknesses. Additional DF imaging shows Cu grains ranging from ~ 20 nm to ~ 100 nm in widths and with morphologies ranging from near equi-axed grains to columnar grains with large length/width ratios. Similar grain structures were observed from all Cu films, with thicknesses ranging from ~ 150 nm to ~ 1180 nm.

Fig. 13 Instrumented indentation of a Cu/Si(001) specimen with a Cu thickness of ~ 1180 nm: (a) indentation modulus versus contact depth, (b) hardness versus contact depth

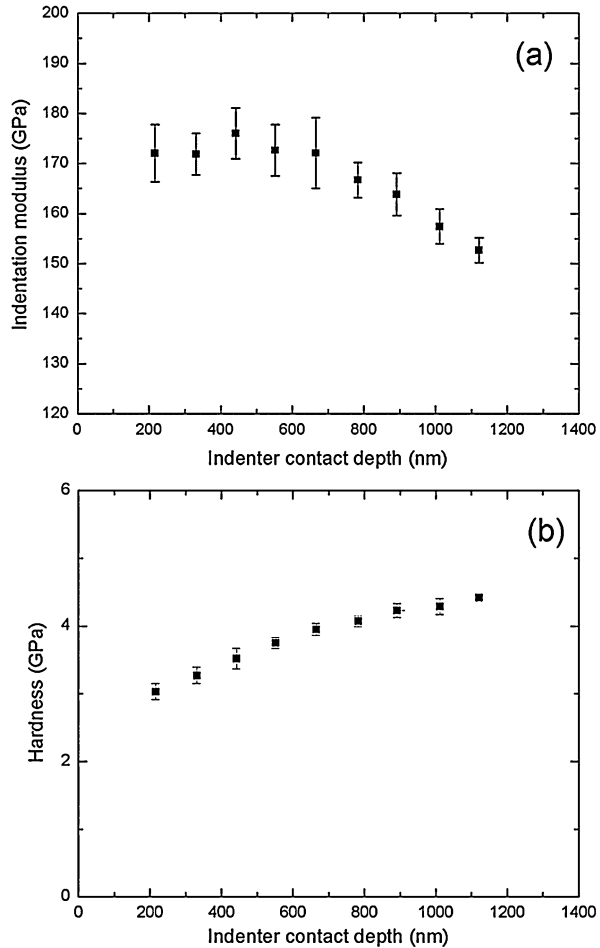


Figure 13 displays indentation results obtained from a Cu/Si(001) specimen, with a Cu film thickness of ~ 1180 nm. Figure 13a shows values of measured indentation modulus, $E_{\text{ind}} = E/(1-\nu^2)$, as a function of the indenter contact depth. Error bars for the modulus value are derived from repeat L - d curves at the same d_{max} value. In the absence of an independent measurement on the Poisson's ratio, only E_{ind} can be obtained from the elastic unloading portion of the indentation curve. Figure 13a shows that E_{ind} increase slightly from ~ 150 GPa at large contact depths to ~ 170 GPa at contact depths between 200 nm and 350 nm.

The elastic stiffness constants for fcc Cu are, respectively, $C_{11} = 168.4$ GPa, $C_{12} = 121.4$ GPa, and $C_{44} = 75.4$ GPa (Kittel 2005), and those for diamond-cubic Si are, respectively, $C_{11} = 165.6$ GPa, $C_{12} = 63.9$ GPa, and $C_{44} = 79.5$ GPa (Hopcroft et al. 2010). The elastic compliance constants for Cu are, respectively, $S_{11} = 14.99$ (TPa) $^{-1}$, $S_{12} = -6.28$ (TPa) $^{-1}$, and $S_{44} = 13.26$ (TPa) $^{-1}$, and those for Si are,

respectively, $S_{11} = 7.69 \text{ (TPa)}^{-1}$, $S_{12} = -2.14 \text{ (TPa)}^{-1}$, and $S_{44} = 12.58 \text{ (TPa)}^{-1}$ (Hopcroft et al. 2010). Young's modulus in the $\langle 001 \rangle$ and $\langle 111 \rangle$ directions of cubic crystals are obtained from the compliance constants as (Nye 2000),

$$E_{100} = \{s_{11}\}^{-1}, \quad (2)$$

$$E_{111} = \left\{ s_{11} - \frac{2}{3} \left[(s_{11} - s_{12}) - \frac{1}{2} s_{44} \right] \right\}^{-1}. \quad (3)$$

For Cu, E_{100} and E_{111} are, respectively, 67 GPa and 191 GPa. For Si, E_{100} and E_{111} are, respectively, 130 GPa and 188 GPa. Furthermore, Poisson's ratio in the $\langle hkl \rangle$ direction of cubic crystals are given by (Zhang et al. 2007)

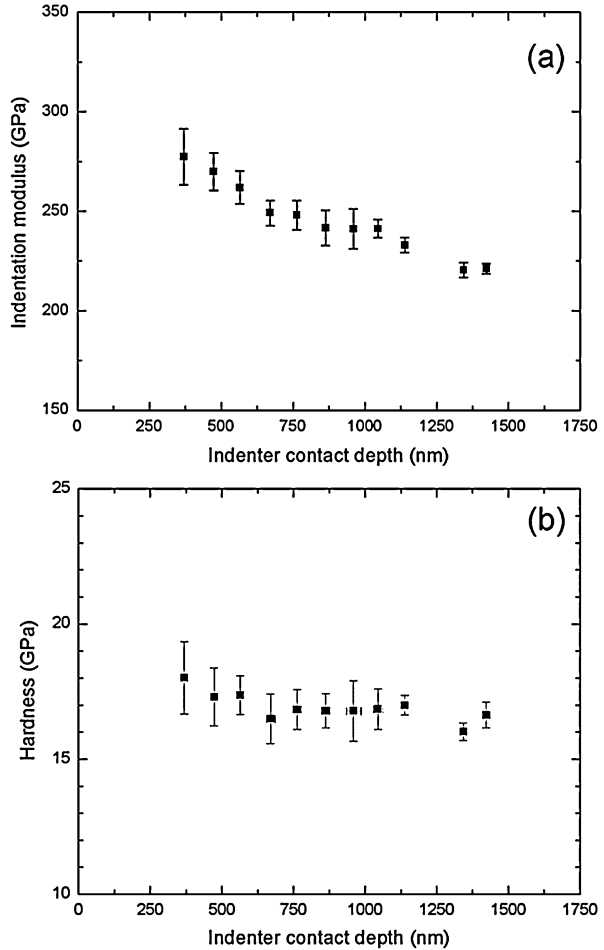
$$\nu_{hkl} = \frac{1}{2} - \frac{E_{hkl}}{2(C_{11} + 2C_{12})}. \quad (4)$$

For Cu, ν_{100} and ν_{111} are, respectively, 0.42 and 0.27. For Si, ν_{100} and ν_{111} are, respectively, 0.28 and 0.18. In the $\langle 001 \rangle$ and $\langle 111 \rangle$ directions, the Cu indentation modulus takes on, respectively, values of 81 GPa and 206 GPa and the Si indentation modulus takes on, respectively, values of 141 GPa and 194 GPa. Based on the XRD results shown in Figs. 8 and 9, the relevant indentation moduli for indentation on the present Cu/Si(001) specimens should, respectively, be $E_{\text{ind}}(111) = 206 \text{ GPa}$ for the Cu film and $E_{\text{ind}}(001) = 141 \text{ GPa}$ for the Si substrate, i.e., the Cu/Si specimen goes from an elastically stiffer film to an elastically more compliant substrate. Thus, the data shown in Fig. 13a are trend-wise consistent with what is expected based on the predominant fiber texture of the Cu films, with Cu $\langle 111 \rangle // \text{Si} \langle 001 \rangle$. Measured E_{ind} value is close to E_{ind} for the Si substrate at large contact depths.

Figure 13b shows values of measured hardness as a function of the indenter contact depth. Error bars for the hardness value are derived from repeat L - d curves at the same d_{max} value. Measured hardness values decrease from above 4 GPa at large contact depths to $\sim 3 \text{ GPa}$ at contact depths between 200 nm and 350 nm. The indentation hardness of Si at room temperature is $\sim 10 \text{ GPa}$, limited by a pressure-induced phase transformation (Vandeperre et al. 2007). The Cu/Si specimen goes from a softer film to a harder substrate. Thus, the observed decrease in hardness with decreasing contact depth is again qualitatively consistent with expectation. Caution should be exercised in taking the measured hardness value of $\sim 3 \text{ GPa}$ at small contact depths as an "intrinsic" hardness of the Cu film, for it is influenced by two extra factors: one being the effect of a harder Si substrate and the other being the expected presence of an indentation size effect. Both effects tend to elevate measured hardness values. For these reasons, indentations on Cu/Si specimens with smaller Cu film thicknesses were not performed.

Figure 14 displays indentation results obtained from a $6.2 \text{ }\mu\text{m}$ thick polycrystalline CrN film deposited on Si(001), over a range of indenter contact depth from $\sim 1400 \text{ nm}$ to $\sim 370 \text{ nm}$. Figure 14a shows that the indentation modulus increases

Fig. 14 Instrumented indentation of a CrN/Si(001) specimen with a CrN thickness of $\sim 6.2\mu\text{m}$: (a) indentation modulus versus contact depth, (b) hardness versus contact depth



monotonically with decreasing indenter contact depth, to ~ 275 GPa at the smallest contact depth of ~ 370 nm. The observed trend is consistent with the expectation for an elastically stiffer film on an elastically more compliance substrate, where measured indentation modulus decreases monotonically with increasing indenter contact depth (Meng and Eesley 1995). A simple extrapolation of the data shown in Fig. 14a suggests that the indentation modulus of the present CrN film is ~ 300 GPa. Figure 14b shows that measured hardness values in the same contact depth range exhibit only a modest change, from ~ 16 GPa to ~ 18 GPa. Noting that values of indenter contact depth normalized to the CrN film thickness range from 0.06 to 0.23, a simple average of all hardness values shown in Fig. 14b yield a hardness value of ~ 17 GPa, a value believed to be representative of the presently deposited CrN films.

Values of Young's modulus for CrN as reported in the literature vary widely, from 190 GPa to above 400 GPa (Chen et al. 2004; Holleck 1986). Moduli of sputter deposited CrN films are often reported to be between 200 GPa and 250 GPa, well below 400 GPa (Sue et al. 1994). A CrN indentation modulus of ~ 300 GPa is consistent with previous reported Young's modulus values of 270–300 GPa on sputter deposited CrN films (Thulasi Raman et al. 2012; Lin et al. 2012). A wide range of CrN hardness values have been reported in the literature, from 10 GPa to above 30 GPa (Zhang et al. 2008). Hardness of CrN was also reported to depend strongly on the level of residual stress (Mayrhofer et al. 2001). The presently measured hardness for CrN of ~ 17 GPa is consistent with the moderate level of residual compressive stress existing within the CrN layer, obtained from estimates based on glancing incidence XRD measurements data shown in Fig. 7.

Micropillar Fabrication and Axial Compression Measurements on Micropillars Containing Coating/Adhesion-Layer/Substrate Interfacial Regions

Fabrication of monolithic cylindrical micro-/nano-pillars using scripted FIB milling has been well documented (Uchic et al. 2004, 2009). To obtain the mechanical response of the interfacial regions in coating/adhesion-interlayer/substrate systems, two groups of micropillar specimens were fabricated: one with interfaces at a 45° inclination to the pillar axis and the other with interfaces perpendicular to the pillar axis (Chen et al. 2014b; Mu et al. 2014). Figure 15 shows a progressive series of SE images obtained during FIB fabrication of CrN/Cu/Si micropillars, in which the interfacial regions are inclined at 45° with respect to the pillar axis. The vapor-deposited CrN/Cu/Si specimens were first mounted at a 45° angle, mechanically polished to reveal the interfacial region, then FIB milled. Figure 15a–c shows the initial stages of top-down annular FIB milling to define the specimen region containing the interfaces and progressively thin down the pillar region. Figure 15d shows a taper-free micropillar resulting from scripted FIB milling with the pillar shown in Fig. 15c turning in a lathe-like motion and the ion beam incident from the side of the pillar. It is evident that this pillar contains the CrN/Cu/Si interfacial regions at a 45° inclination. Figure 15e shows such an array of taper-free micropillars. Axial compression testing of such micropillar arrays allows repeated and independent testing of the coating/substrate interfacial regions in a combined compression/shear loading. Similar top-down and scripted FIB milling was used to fabricate micropillars with interfacial regions perpendicular to the pillar axis. Figure 16 shows close-up SE images of typical CrN/Cu/Si(001) micropillars with interfaces inclined at 45° and 90° with respect to the pillar axes. Additional information on pillar fabrication was provided elsewhere (Chen et al. 2014b; Mu et al. 2014).

In situ axial compression of micropillars was conducted on a NanoMechanics Inc. NanoFlip[®] instrumented micro-/nano-mechanical testing device placed within the FEI Quanta3D FIB instrument. A custom-made, flat-ended, $\sim 10 \mu\text{m} \times 10 \mu\text{m}$ diamond punch was used for pillar compression. The NanoFlip device, interfaced with the Quanta3D instrument, enabled simultaneous acquisition of the total load force on the indenter, P , the total punch displacement, Δ , and a video of SE images

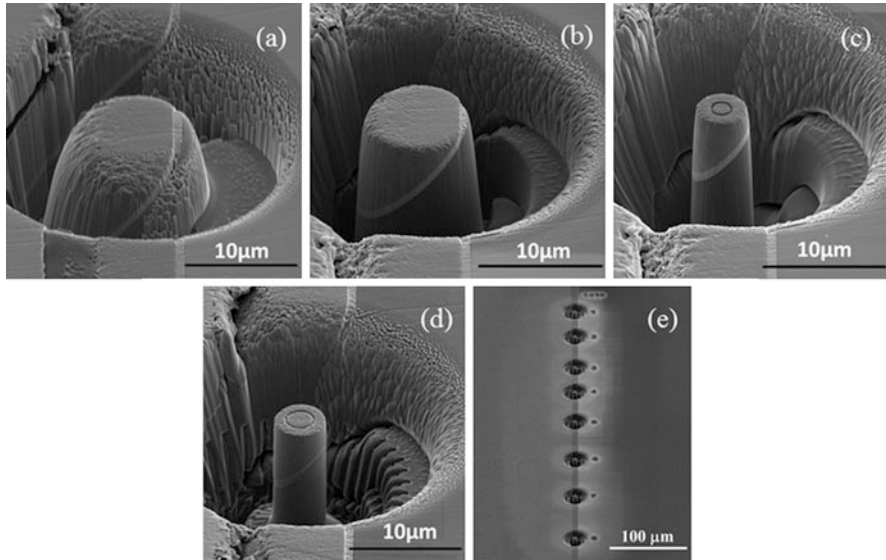


Fig. 15 Fabrication of micropillars containing interfacial regions of coating/adhesion-interlayer/substrate specimens: (a) top-down annular FIB milling after mechanical polishing, (b) defining specimen region with interfaces at a 45° inclination with respect to the pillar axis, (c) progressive thinning down of the pillar region, (d) the final pillar after scripted FIB milling with pillar turning and Ga^+ ion incident from the pillar side, (e) A finished pillar array, enabling repeated testing of the same interfacial region in independent experiments

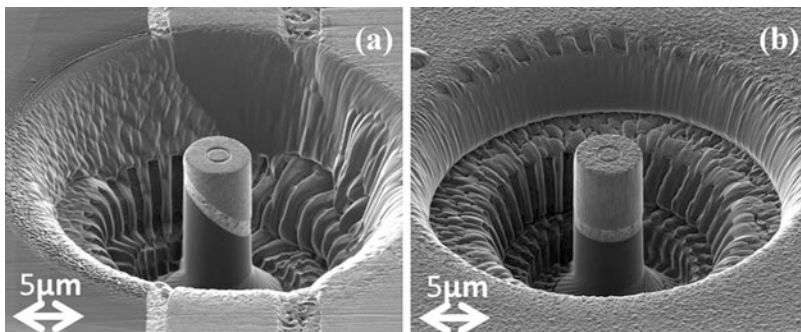


Fig. 16 Typical cylindrical micropillars fabricated by scripted Ga^+ FIB milling: (a) one CrN/Cu/Si(001) micropillar with interfaces inclined at 45° to the pillar axis, (b) one CrN/Cu/Si(001) micropillar with interfaces perpendicular to the pillar axis

of the micropillar under testing, thus allowing specific points on the $P-\Delta$ curve to be linked to specific images of the pillar morphology at that point. The NanoFlip is fundamentally a load-controlled instrument, but is programmed to enable loading to occur in a “displacement-controlled” mode by providing feedback control of loading force according to deviation of sensed punch displacement rate from the set value.

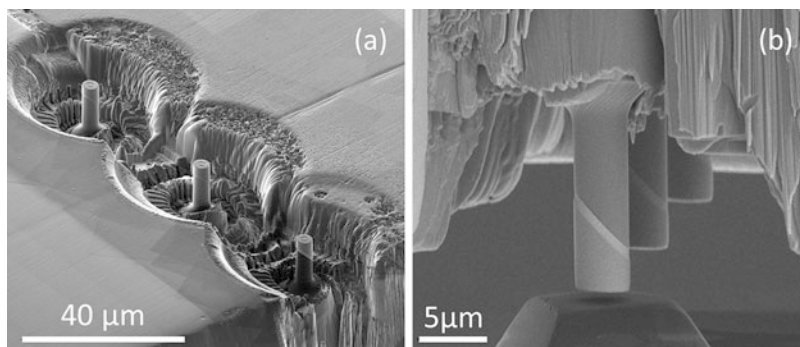


Fig. 17 Instrumented microscale compression testing of micro-pillar specimens in situ an SEM: (a) an array of three micropillars with circular surrounding openings to allow for approach of a diamond indenter. (b) the experimental configuration prior to contact of the diamond indenter with one micropillar

All in situ pillar compression tests were performed under such a displacement-controlled mode, with the target displacement rate set at 10 nm/s. Additional ex situ axial compression testing of pillars was conducted on a NanoIndenter XP instrument with another custom, flat-ended, $\sim 10 \mu\text{m} \times 10 \mu\text{m}$ diamond punch. These ex situ tests were conducted without concurrent observations, and only post test morphological observations of the tested pillars were made after the compression testing had concluded.

Cylindrical micropillars of CrN/Cu/Si(001) and CrN/Ti/Si(001) specimens, 3–5 μm in diameter, were fabricated with the interfacial regions inclined at 45° and 90° with respect to the pillar axis. From one CrN/interlayer/Si specimen, a number of micropillars were fabricated such that repeat axial compression tests can be performed. The typical experimental situation related to in situ axial compression testing of CrN/interlayer/Si(001) micropillars is illustrated in Fig. 17. The FIB milling process yielded cylindrical micropillars, 3–5 μm in diameter and $\sim 10 \mu\text{m}$ in length, resting on a larger Si base. Figure 17a shows an array of three micropillars fabricated from one CrN/Cu/Si(001) specimen, with the Cu interlayers inclined at 45° with respect to the pillar axes. The Cu interlayer thickness for this specimen is ~ 550 nm, measured from SE images of a FIB cross section of the specimen. Finished pillars are taper-free and are subjected to no mechanical contact once the milling process is completed. The circular opening surrounding each pillar, $\sim 30 \mu\text{m}$ in diameter, is sufficiently large to allow the $\sim 10 \mu\text{m} \times 10 \mu\text{m}$ diamond punch to engage the pillar top without contacting the rim of the circular opening. As shown in Fig. 17a, the right most pillar is located near the specimen edge. The rim material bordering the specimen edge, as well as the rim material separating one pillar from another, has been removed by additional FIB milling to allow SE imaging during axial compression of the pillars. The SE image shown in Fig. 17b illustrates the situation immediately before the flat face of the $\sim 10 \mu\text{m} \times 10 \mu\text{m}$ diamond punch is about to engage the top of the right most pillar shown in Fig. 17a, with the two

other pillars visible in the same field of view. The shadow on the flat face of the punch, due to the pillar blocking a portion of secondary electrons emanating from the punch flat face, serves to locate the pillar in the center of the punch face during the engagement process. Similar pillar morphologies and in situ pillar compression test configurations hold for CrN/Ti/Si micropillars with interfaces inclined at either 45° or 90° with respect to the pillar axes.

Figure 18a shows the raw P - Δ curve obtained from in situ axial compression of the right most CrN/Cu/Si(001) micropillar shown in Fig. 17a, with a Cu interlayer thickness of ~ 550 nm and the interfaces inclined at 45° with respect to the pillar axis. Figure 18a shows an initial rapid increase in P with increasing Δ until a critical load value, P_c , is reached. Once P_c is reached, Δ increases with little further increase in P , exhibiting a stable load plateau. This load plateau, however, does not extend indefinitely. Rather, large and discontinuous punch displacement excursions occur as Δ increases beyond a point. Figure 18b-d shows single frames of the SE imaging video of the same pillar at various stages during the axial compression experiment, corresponding to the points 1, 2, and 3 identified on the P - Δ curve shown in Fig. 18a, with total punch displacements at ~ 115 nm, ~ 370 nm, and ~ 440 nm, respectively. It is apparent from Fig. 18a, b that little or no deformation of the Cu interlayer is observed at point 1 on the rapid load rise portion of the P - Δ curve. Once the critical load P_c is reached and the punch displacement progresses to point 2 on the load plateau, the entire Cu interlayer undergoes a shear deformation. The upward displacement of the punch is accommodated by a rigid upward shift of the top CrN portion of the pillar and is in turn accommodated by a shear deformation of the entire Cu interlayer. The CrN top portion of the pillar shifts rigidly upward and to the left in Fig. 18c. This mode of deformation continues as the punch displacement progresses further on the load plateau to point 3, and the entire Cu interlayers undergoes an increased amount of shear deformation, as evident from Fig. 18d. The large and discontinuous punch displacement excursion observed, to beyond the total punch displacement range of 800 nm displayed in Fig. 18a, is concomitant with a catastrophic shear failure. As shown in Fig. 18e, this dynamic shear-off event causes the diamond punch to quickly move upward in an uncontrolled manner, impacting the bottom Si portion of the pillar and destroying it. The entire dynamic shear-off event occurs suddenly, within the time span for collecting one single SE image frame, thus leaving no image of the pillar during any intermediate stage of this dynamic shear-off. The SE image in Fig. 18f illustrates separate posttest morphological examinations of CrN/Cu/Si pillars subjected to ex situ axial compression, after the large punch displacement excursion had occurred. It is clear that shear loading on the interfacial region, due to the axial compression load, caused a shear failure of the interfacial region, resulting in the top CrN portion of the pillar shifting rigidly with respect to the Si bottom portion. Spot mode EDS spectra collected from the top surface of the exposed bottom pillar portion and the bottom surface of the exposed top pillar portion both show the presence of Cu signal, indicating that the dynamic shear-off occurred within the Cu interlayer. The SE image shown in Fig. 18f and multiple similar ones, not shown, further indicate that the final dynamic shear-off occurred near the Cu/CrN interface, instead of in the

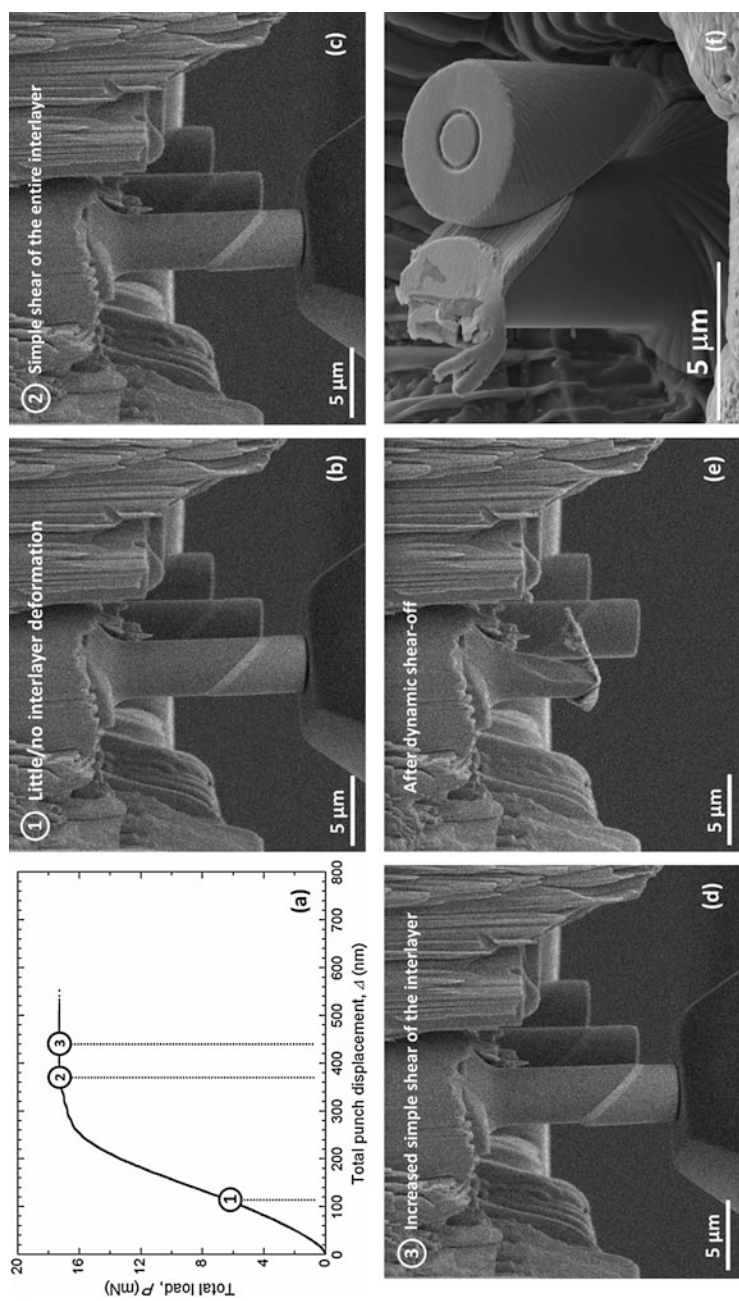


Fig. 18 In situ compression testing of one CrN/Cu/Si(001) micro-pillar with 45° inclined interfaces: (a) the raw $P-\Delta$ curve, points 1, 2, and 3 correspond, respectively, to Δ values of ~ 115 nm, ~ 370 nm, and ~ 440 nm, (b) the SE image of the pillar at $\Delta \sim 115$ nm (point 1) (c) the SE image of the pillar at $\Delta \sim 370$ nm (point 2), (d) the SE image of the pillar at $\Delta \sim 440$ nm (point 3), (e) the SE image of the pillar after occurrence of the large and discontinuous displacement excursion, (f) a separate posttest SE image of another CrN/Cu/Si pillar after the dynamic shear-off event occurred, illustrating the consequence of an interfacial shear instability

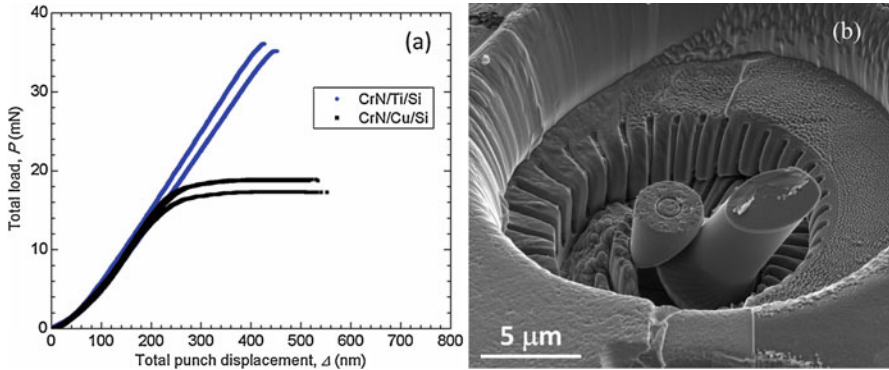


Fig. 19 In situ compression testing of CrN/interlayer/Si(001) micropillars with 45° inclined interfaces: (a) raw P - Δ curves obtained from testing pillars with Cu and Ti interlayers, (b) a posttest SE image of a pillar with a Ti interlayer, after the occurrence of the unstable dynamic shear-off event

middle of the Cu interlayer. Correlating the P - Δ curve, the concurrent SE imaging during in situ compression testing, and additional posttest SE images obtained from ex situ compression tested pillars, it is concluded that axial compression of CrN/Cu/Si micropillars with 45° inclined interfaces results in shear failure of the interfacial region and that this shear failure occurs in two stages: first by a shear deformation of the entire metal interlayer, then followed by an unstable dynamic shear-off near the Cu/CrN interface.

Figure 19a shows a collection of P - Δ curves obtained from in situ compression testing of CrN/Cu/Si and CrN/Ti/Si micropillars. The three P - Δ curves for CrN/Cu/Si were collected from testing of the three micropillars shown in Fig. 17a. Qualitatively similar P - Δ curves were observed from all three pillars: with an initially rapid rise in P with increasing Δ , followed by a stable and extended load plateau, and finished with large and discontinuous punch displacement excursions concomitant with the final dynamic shear-off. The two P - Δ curves for CrN/Ti/Si were collected from compression testing of two micropillars fabricated from one CrN/Ti/Si(001) specimen. The CrN/Ti/Si pillars are again $\sim 4 \mu\text{m}$ in diameter and $\sim 10 \mu\text{m}$ in length, resting on a larger Si base. The Ti interlayer thickness is $\sim 340 \text{ nm}$ and the interfaces are 45° inclined with respect to the pillar axis. These P - Δ curves exhibit an initially rapid rise in P with increasing Δ , but show very limited load plateaus when a critical load P_c is reached, followed by large and discontinuous punch displacement excursions concomitant with the final dynamic shear-off.

Comparing P - Δ curves shown in Fig. 19a, the differences in the response of CrN/Ti/Si pillars subjected to axial compression as compared to that of CrN/Cu/Si pillars are apparent. First, the critical load values are significantly different: P_c values are $\sim 18 \text{ mN}$ and $\sim 36 \text{ mN}$ for CrN/Cu/Si and CrN/Ti/Si pillars, respectively. Second, the extent of the load plateau is much more limited for the CrN/Ti/Si pillars

as compared to the extended and stable load plateaus observed for the CrN/Cu/Si pillars. The absence of an extended and stable load plateau during axial compression of CrN/Ti/Si pillars means that very soon after the axial load reaches P_c , dynamic shear-off occurs. The SE image videos during pillar compression testing of CrN/Ti/Si pillars thus contain no frame showing either significant deformation of the Ti interlayer in shear or any intermediate stage of the final dynamic shear-off. Figure 19b shows posttest morphological examination of one CrN/Ti/Si pillar subjected to ex situ axial compression, after the final large punch displacement excursions had occurred. In this case, the punch lurching forward knocked the bottom pillar portion off its base. The top CrN portion of the pillar shifted as a whole with respect to the Si bottom portion along the 45° inclination as a result of the shear loading on the interfacial region. Spot mode EDS analysis showed the presence of Ti signals on both the top surface of the bottom pillar portion and the bottom surface of the top pillar portion. Similar to Fig. 18f, examination of the image shown in Fig. 19b shows that the final dynamic shear-off occurred near the Ti/CrN interface, instead of in the middle of the Ti interlayer.

Data and observations shown in Figs. 18 and 19 indicate that axial compression of CrN/interlayer/Si micropillars with 45° inclined interfaces led to shear failures of the interfacial regions in two stages: an initial shear deformation of the entire metal interlayer, followed by a sudden dynamic shear-off close to the metal/CrN interface. The main difference between the CrN/Cu/Si system and the CrN/Ti/Si system subjected to shear loading on the interfacial region is the existence of an extended and stable load plateau in the former system and a limited load plateau in the latter system. For CrN/Cu/Si, the entire Cu interlayer is subjected to a shear deformation to large strains before the final dynamic shear-off occurs close to the Cu/CrN interface. For CrN/Ti/Si, the final dynamic shear-off close to the Ti/CrN interface occurs soon after the Ti interlayers are subjected to a shear deformation.

Because shear failure of the interfacial regions initiates with a shear deformation of the entire metal interlayer, the critical load P_c measured experimentally from axial compression of CrN/interlayer/Si micropillars with 45° inclined interfaces yields a measure of the average shear stress τ_c necessary to initiate shear failure of the interfacial region,

$$\tau_c = P_c / (\pi D^2/4) / 2, \quad (5)$$

where D is the pillar diameter. Figure 20 summarizes results of ex situ and in situ axial compression testing on CrN/Cr/Si(001) and CrN/Ti/Si(001) micropillars with 45° inclined interfaces by plotting values of τ_c against the metal interlayer thickness h . Separate data points at the same h value denote repeat measurements on separate pillars. From data presented in Fig. 20, several points should be noted. First, τ_c values measured from ex situ and in situ compression tests, conducted with two separate instruments on the same specimens, are consistent, indicative of the fidelity of the present set of measurements. Second, the value of τ_c increases with decreasing h , significantly for Cu interlayers, somewhat for Ti interlayers although the data scatter is large in this case. An increase in τ_c with decreasing h is

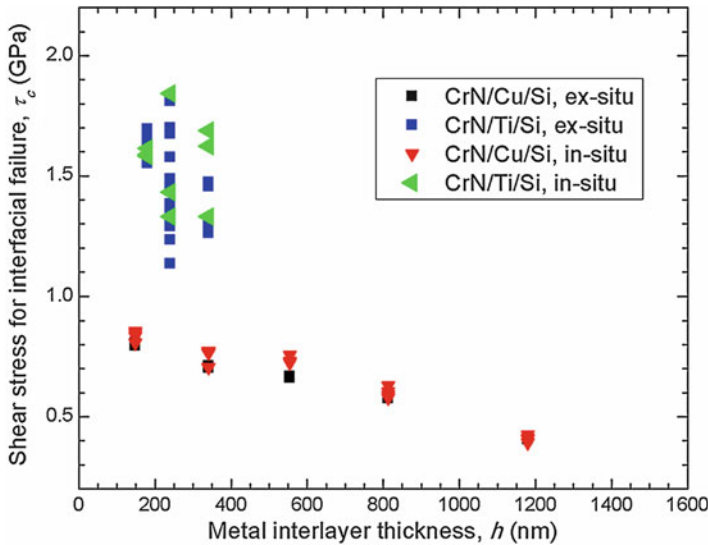


Fig. 20 Average shear stress to initiate shear failure of the interfacial regions of CrN/interlayer/Si specimens plotted versus the thickness of the metal interlayer

qualitatively consistent with strain gradient plasticity model predications, according to which confinement of the metal interlayer in between two elastic-brittle solids offers one reason for the observed increase in interlayer flow stress (Mu et al. 2014). Third, τ_c depends significantly on the metal interlayer: the shear failure stress for Ti interlayers exceeds that for Cu interlayers by about a factor of two, significantly outside the data scatter band even with the large scatter in the Ti case.

Clues to the mechanism through which the final dynamic shear-off occurs in the two CrN/interlayer/Si systems are offered from observations made during in situ axial compression testing of CrN/Cu/Si(001) micropillars with interfaces 90° inclined with respect to the pillar axes. Figure 21a shows a raw $P-\Delta$ curve obtained from in situ axial compression of one CrN/Cu/Si(001) micropillar, with a Cu interlayer thickness of ~ 1180 nm and the interfaces inclined at 90° with respect to the pillar axis. Figure 21a shows an initial rapid increase in P with increasing Δ until a critical load P_c is reached. Once P_c is reached, the $P-\Delta$ curve exhibits a plateau region in which P increases only slowly with increasing Δ . Beyond this plateau region, further increases in Δ bring again significant increase in P . Figure 21b–e shows single frames of the SE imaging video of the same pillar at various stages during the axial compression experiment, corresponding to the points 1, 2, 3, and 4 identified on the $P-\Delta$ curve shown in Fig. 21a, with total punch displacements at ~ 225 nm, ~ 425 nm, ~ 525 nm, and ~ 1200 nm, respectively. It is apparent from Fig. 21b that little or no deformation of the Cu interlayer is observed at point 1 on the rapid load rise portion of the $P-\Delta$ curve. Once the critical load P_c is reached and the punch displacement progresses to point 2 on the load plateau, Fig. 21c shows

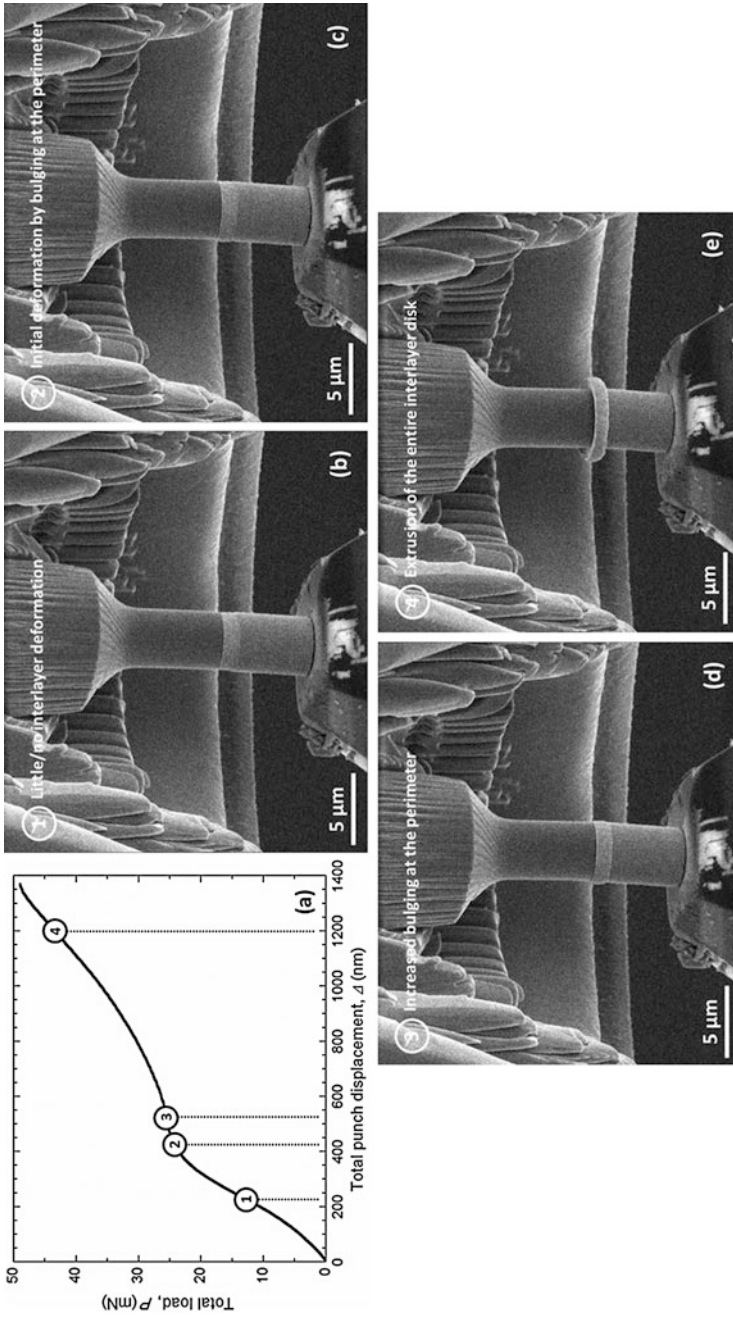


Fig. 21 In situ compression testing of one CrN/Cu/Si(001) micropillar with 90° inclined interfaces: (a) the raw P - Δ curve, points 1, 2, 3, and 4 correspond, respectively, to Δ values of ~ 225 nm, ~ 425 nm, ~ 525 nm, and ~ 1200 nm, (b) the SE image of the pillar at $\Delta \sim 225$ nm (point 1), (c) the SE image of the pillar at $\Delta \sim 425$ nm (point 2), (d) the SE image of the pillar at $\Delta \sim 525$ nm (point 3), (e) the SE image of the pillar at $\Delta \sim 1200$ nm (point 4)

that deformation of the Cu interlayer occurs by bulging out at the perimeter from the center of the interlayer, with the diameters of the Cu interlayer at the Cu side of the Cu/ceramic interfaces appearing not to change. This mode of deformation continues as the punch displacement progresses further on the load plateau to point 3, as shown in Fig. 21d. The entire Cu interlayer undergoes an increased amount of bulging out at the perimeter from the center of the interlayer, still leaving the diameters of the Cu interlayer at the Cu side of the interfaces unchanged. As the punch displacement progresses to beyond the load plateau at point 4 on the $P-\Delta$ curve, the entire Cu interlayer is extruded out from the perimeter with its diameter significantly larger than the original pillar diameter, as shown in Fig. 21e.

The in situ compression testing of CrN/Cu/Si micropillars with 90° inclined interfaces offers one insight to how the metal interlayers confined between elastic-rigid materials, CrN and Si in the present case, deform under load. As shown in Fig. 22a, the morphology of the Cu interlayer during an early elasto-plastic stage of deformation is illustrated with an actual SE image of a deformed Cu interlayer. In this posttest examination, Ga^+ FIB milling was used to remove a small section of the pillar after compression loading to reveal better the morphology of the deformed Cu interlayer. At this early deformation stage, plastic deformation of the interlayer occurs by bulging out at the perimeter in the center, while the metal/ceramic interfaces appear to be “locked,” with no relative displacement between the metal and ceramic side of the interface. Figure 22b shows two analogous SE images of a deformed Cu interlayer at a later deformation stage. The interlayer deformation leads to the entire interlayer disk being extruded out in an axisymmetric manner. Here the metal/ceramic interfaces appear to be “unlocked,” with significant relative displacement between the metal and ceramic side of the interface, thus allowing extrusion of the entire interlayer disk to occur. The sliding marks, visible on the top surface of the extruded Cu interlayer disk, offer clear evidence that relative displacement occurred between the Cu side and the CrN side of the Cu/CrN interface. The FIB milled cut-away view in Fig. 22c shows that the thickness of the Cu interlayer disk is the largest at the outer rim, further indicating that extrusion of the Cu interlayer disk occurred after the interface became “unlocked.” Unfortunately, plastic flow stresses of thin Ti layers are too high to allow similar observations to be made in the cases of CrN/Ti/Si pillars: the Si bottom portion of the pillars fail in compression before perceptible deformation of the Ti interlayer occurs in the confined normal compression geometry. Nevertheless, we suggest that a similar transition occurs for both CrN/Cu/Si and CrN/Ti/Si interfacial regions, i.e., the interfaces go from being “locked” to being “unlocked” as the metal interlayers confined between CrN and Si are stressed to increasingly higher levels. We further suggest that this interfacial “unlocking” is physically related to the final dynamic shear-off observed in both systems, as illustrated in Figs. 18 and 19. To our knowledge, both the observation of an interfacial instability under shear loading and the suggestion that this shear instability is linked to a transition from an interfacial “locking” to an interfacial “unlocking” are new and demonstrate the potential of in situ instrumented microscale mechanical testing and its application to studying the mechanical integrity of metal/ceramic interfaces.

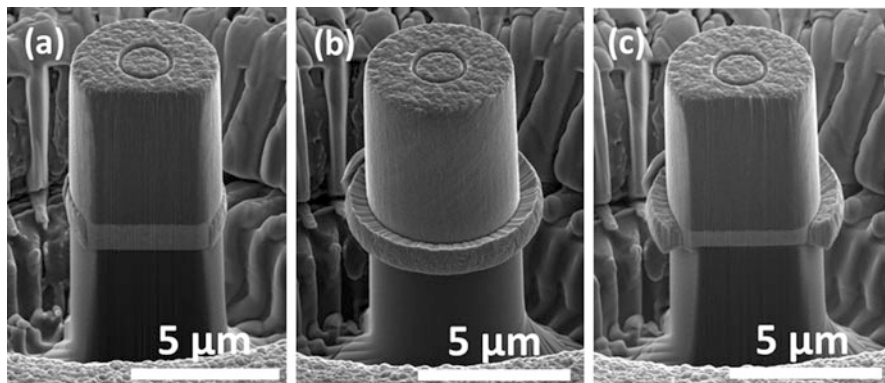


Fig. 22 Posttest images of deformed Cu interlayer: (a) in the early elasto-plastic stage, (b) in the later elasto-plastic stage

Mechanistic Understanding on the Shear Instability of Metal/Ceramic Interface Provided by Simulations

In this subsection, the structure and mechanical response of the Ti/TiN interfaces are investigated using MD/DFT simulations to shed light on the relevant properties of a general metal/ceramic interface. Part of the reasons for this choice is the availability of the semi-empirical potentials available for metal/ceramic systems (Zhang et al. 2017; Yang et al. 2017; Kim and Lee 2008a). At room temperature, the equilibrium lattice structure of Ti is HCP while that for TiN is of NaCl type. The preferred orientation relation (OR) for the Ti/TiN system is $\langle 0001 \rangle_{Ti} // \langle 111 \rangle_{TiN}$ and $\langle 11\bar{2}0 \rangle_{Ti} // \langle 110 \rangle_{TiN}$ on the interface plane: $(0001)_{Ti} // (111)_{TiN}$ (Yang et al. 2017; Sant et al. 2000). Another reason for the choice of the Ti/TiN system is the similarity between TiN and CrN – material used in our experiments (section “[Synthesis of Ceramic–Coating/Adhesion–Interlayer/Substrate Systems and Structural/Mechanical Characterization of Coating/Substrate Interfacial Regions](#)”). Both nitrides have identical lattice type and very close lattice parameters ($a_{TiN} = 4.242 \text{ \AA}$ vs. $a_{CrN} = 4.149 \text{ \AA}$) (Toth 1971; Kim and Lee 2008b). The structures of the Ti/TiN, Ti/CrN, and Cu/CrN interfaces are also expected to be similar, since the hexagonal atomic structure on (0001) Ti planes is also similar to the (111) Cu planes (Shao et al. 2013; Wang et al. 2014). Therefore, understandings gathered from the numerical modeling efforts herein, namely, the structure and properties of Ti/TiN interfaces, should have significant implications on other metal/ceramic interfacial systems, e.g., Ti/CrN, Cu/CrN (Shao et al. 2015; Wang et al. 2008, 2014; Zhang et al. 2013; Chen et al. 2017).

As an outline, in section “[The Structure of the Misfit Dislocation Network and the Coherent Regions of the Ti–TiN Interfaces](#),” following the methodology proposed by Hirth and Lothe (1982), the structure of the Ti–TiN interface is analyzed through the “structural relaxation approach.” As will be shown, the structural details of the

semi-coherent interface are dependent on the intrinsic properties of the coherent Ti-TiN interfacial structures. Hence, in section “The Location of the Misfit Dislocation Network,” the energetic characteristics of the coherent structures evaluated by DFT calculations are presented. Finally, in section “Intrinsic Energetic Characteristics of Coherent Ti/TiN Interfacial Structures,” the mechanical response of the Ti-TiN interfaces to external shear loading is discussed.

The Structure of the Misfit Dislocation Network and the Coherent Regions of the Ti-TiN Interfaces

Two specific ORs are permitted by the general OR given above, namely:

- OR 1: $x//[1\bar{1}00]_{Ti}//[11\bar{2}]_{TiN}$, $y//[0001]_{Ti}//[111]_{TiN}$, and $z//[11\bar{2}0]_{Ti}//[1\bar{1}0]_{TiN}$,
- OR 2: $x//[1\bar{1}00]_{Ti}//[11\bar{2}]_{TiN}$, $y//[0001]_{Ti}//[111]_{TiN}$, and $z//[1\bar{1}20]_{Ti}//[1\bar{1}0]_{TiN}$.

On the interface, the $(111)_{TiN}$ plane and the $(0001)_{Ti}$ plane have the identical hexagonal structure, with lattice match of 1.84%. Due to this mismatch, a number of distinct coherent structures may exist on the unrelaxed Ti/TiN interfaces of the above two ORs, as shown by the plan views of provided in Fig. 23a, b. Such coherent structures include near-FCC, near-HCP, near-Overlap structures and the region separating the former three, based on the relative positions of Ti atoms straddling on both size of the interface, which are shown in the lower portions of Fig. 23. Even though the Ti atoms may have taken the identical relative positions in

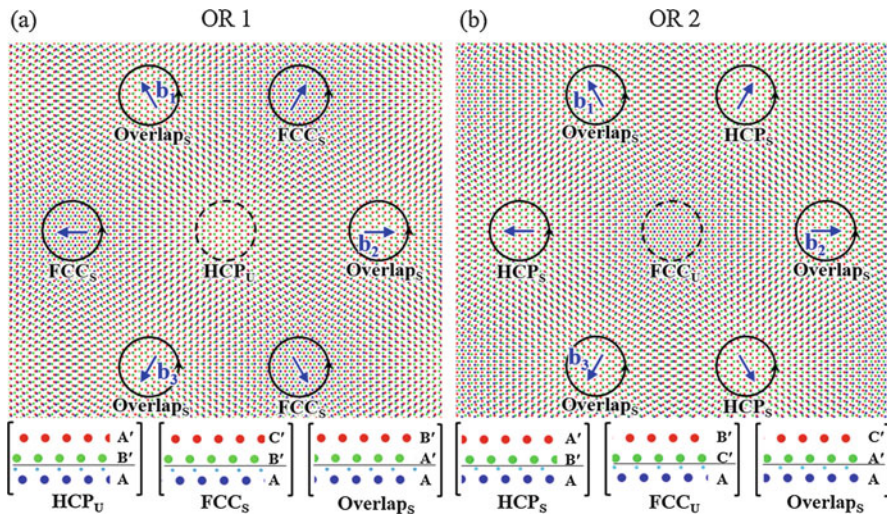


Fig. 23 Structure of the “unrelaxed” Ti/TiN interface prior to formation of the MDN at the chemical interface. The plan views of the interface are provided on the top portions of the figure, and the local atomic stacking is provided on the lower portion. The N atoms are omitted in the plan view for clarity and are shown as smaller spheres in the stacking structures. The atoms are colored according to their vertical positions. The primed notations, i.e., A’B’A’, etc., is used for Ti layer to distinguish from stacking in TiN layer. The thin solid lines denote the position of the chemical interface

the respective FCC and HCP regions in the Fig. 23a, b, the structures are clearly not identical, due to the presence of the N atoms. DFT calculations show that (details in the next sections) the FCC and HCP structures in the two ORs have different energy levels. For instance, the FCC structure in OR 1 has the lowest excess energy; therefore, the relative positions of the Ti and N atoms specified in this structure are referred to as the FCC_S structure. On the other hand, the FCC in OR 2, in which the Ti atoms on the Ti side overlap with the N atoms, corresponds to a higher excess energy. According to DFT calculations, this structure is unstable and is thus named FCC_U. Same naming convention and energy characteristics apply for the HCP structures. Unlike the FCC and HCP structures, the Overlap structures are metastable and are identical in both ORs.

Following the approach proposed by Hirth and Lothe (1982), the regions with stable and metastable atomic structure can be regarded as being enclosed by Shockley partial dislocation loops (black solid circles shown in Fig. 23). The loops have Burgers vectors: b_1 , b_2 , b_3 , $-b_1$, $-b_2$, and $-b_3$ ($b_1 = \frac{a}{3} [0\bar{1}10]$, $b_2 = \frac{a}{3} [10\bar{1}0]$, and $b_3 = \frac{a}{3} [\bar{1}100]$). The arrows on the loops mark the line senses of the loops. Upon relaxation, driven by the minimization of the interface excess energy, the regions of FCC_S, HCP_S, and Overlap_S expand. As a result, larger patches of coherent regions form associated with the in-plane biaxial straining of the Ti and TiN lattices. When the adjacent loops reach close proximity with each other, they react and form the misfit dislocation network (MDN) (schematically shown in Fig. 24, where partial dislocation loops, misfit dislocation lines and nodes are represented in light blue, brown, and green). The nodes (intersections of dislocations) of the MDN originate from the contraction of regions with unstable structures. Dislocation lines in the MDN are also Shockley partials with Burgers vectors b_1 , b_2 , and b_3 . The thermodynamics of the interface relaxation is consistent with the perspective of the classical Peierls-Nabarro dislocation model (Peierls 1940; Nabarro 1947) and is characterized by two competing factors, namely: (1) the reduction of the chemical potential energy (excess interface energy of the coherent structures) and (2) the increase in the core strain energy in the MDN. The former factor is governed by the generalized stacking fault energy (GSFE) of the interface, while the latter is governed by the normal elastic constant of the adjoining crystals. In effect, the core width of the MDN inversely relates to the magnitude of interfaces' GSFE profile (Shao et al. 2014), for instance, higher amplitude in GSFE corresponds to narrower cores.

Two distinct relaxed interface structures exist for the two ORs (Figs. 25a, b). Interface with OR 1 comprises FCC_S (green atoms) and Overlap_S (white atoms) regions, while the interface with OR2 comprises HCP_S (red atoms) and Overlap_S (white atoms) regions. It is interesting to note that the dislocation lines of the MDN (on the edge of the regions in Fig. 25a, b) on the chemical interface are jagged. This is consistent with the observation made by Yang et al. in their recent work (Yang et al. 2017). Accordingly, the blue solid lines in Fig. 25a, b simply mark the neutral positions of the dislocations. Therefore, the net character on dislocation lines is edge, even though the local character of the partial misfit dislocations may be

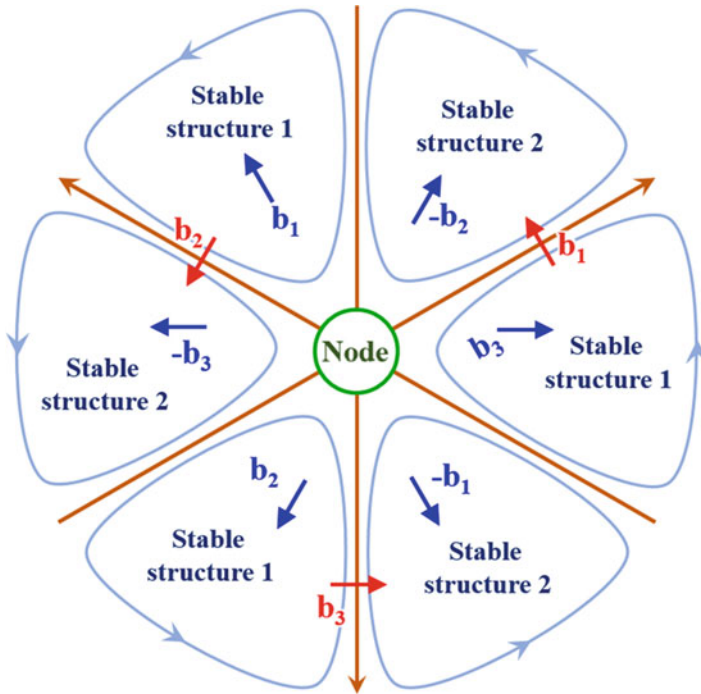


Fig. 24 Schematic of the expansion of and reaction between the partial dislocation loops which encloses the coherent stable structures

mixed. In addition, from Fig. 25a, b, one can see that both FCC_S (green) and HCP_S (red) regions occupy greater area than the $Overlap_S$ (white) regions, also the area of the HCP_S regions is greater than the area of the FCC_S regions. This implies that the excess specific energies of the coherent structures have the following relation, $E_{Overlap_S} > E_{FCC_S} > E_{HCP_S}$. Our DFT calculations confirm this prediction, as will be shown in the next sections.

The Location of the Misfit Dislocation Network

The energetically preferred location of the MDN, in general, should not be at the chemical interface. The reasoning is given as follows:

1. The elastic line energy (per unit length) of dislocations is proportional to the shear modulus of the material, i.e., $E/L \propto Gb^2$. A dislocation of the same burgers vector in ceramics would have significantly higher energy than one in metal. Therefore, from the thermodynamics standpoint, the MDN would have significantly higher energy when it is located at the Ti/TiN chemical interface compared to when it is in Ti.
2. The stress field of misfit dislocations decays exponentially as a function of distance from the plane of MDN (Hirth and Lothe 1982; Shao et al. 2018).

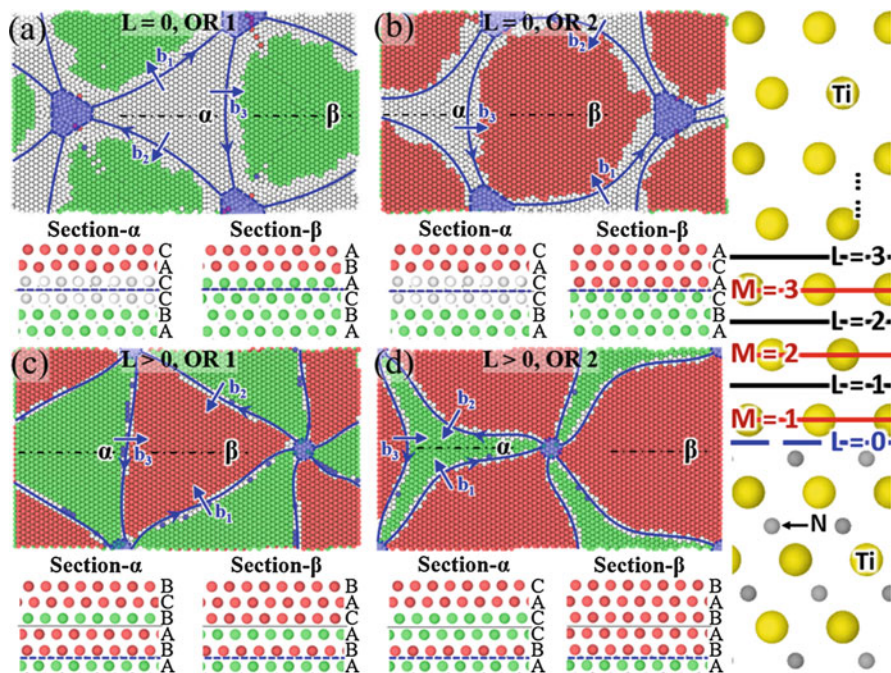


Fig. 25 Typical structures of the MDNs corresponding to different locations, i.e., different L numbers as indicated on the right. (a) and (b) show Type I and Type II structures for $L = 0$, (c) and (d) show Type I and Type II structures for $L > 0$. Atoms are colored by common neighbor analysis performed on Ti atoms. Atoms within a local HCP structure are colored red, and those within a local FCC structure are colored green. Atoms that are not in HCP, FCC, BCC, or icosahedral structures are colored white. Misfit dislocation lines and nodes are denoted by blue lines and blue shades

Consequent, the strain energy of the MDN is expected to decrease sharply as it moves away from the chemical interface into Ti.

3. As an MDN is moving away from the chemical interface into Ti, a layer of Ti with in-plane lattice constant coherent to TiN is created. The thickness of this layer is identical to the distance between the MDN to the chemical interface. The total strain energy of this layer therefore increases linearly with its thickness.
4. The true interface configuration depends on the total free energy of the entire interface region (three dimensional entity), which includes the chemical potential energy at the chemical interface, the elastic energy of the misfit dislocation network, and the elastic energy of the coherent Ti atomic layers.

Therefore, it is expected that the MDN prefers to reside inside Ti but very close to the chemical interface. This is confirmed by Yadav et al. (2015) in their recent work, where they found that the interface energy reaches minimum when the MDN is in Al and 2 atomic monolayers away from the chemical interface.

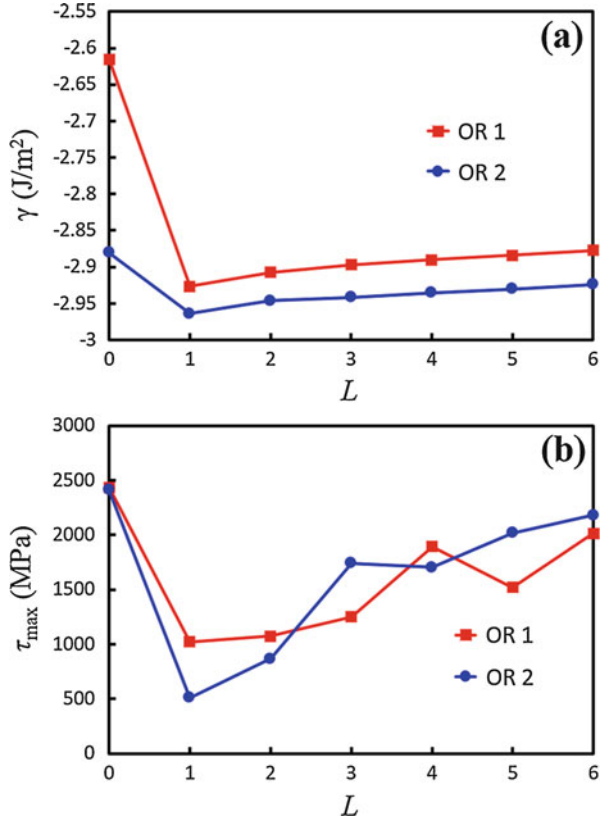
In this section, we present our prior findings, obtained from molecular dynamics simulations, on the variation in the structure as well as the energy of the interface as a function of the position of the MDN in Ti. An L parameter is used here to indicate the position of the MDN. As shown on the right hand side of Fig. 25, the MDN resides either on the chemical interface (blue dashed line), which is in between the atomic monolayers of Ti and N ($L = 0$), or on one of the (0001) slip planes in Ti ($L > 0$, black solid lines). The (0001) slip planes exist in between two adjacent (0001) atomic monolayers of Ti. When the MDN is inside Ti and has n Ti atomic monolayers between the chemical interface, $L = n$. The parameter M indicates the position of (0001) Ti atomic monolayers (red solid lines) near the interface, with $M = 1$ denoting the Ti atomic layer most adjacent to the chemical interface. The structure of the Ti/TiN interface for $L = 0$ has been discussed in detail above. Similar to the case of $L = 0$, the coherent regions are separated by the MDN for $L > 0$. For OR 1, the coherent regions have the structures that comprise single layer of atoms surrounded by local FCC structure (green atoms) and a single layer of atoms surrounded by HCP structure (red atoms). The structures are commonly referred to as the type 1 intrinsic stacking fault (I_1) in HCP (Benoit et al. 2013) and are energetically equivalent. Thus, the area fraction of the coherent regions is the same (see the green and red regions in Fig. 25c). For OR 2, the coherent regions include a higher energy structure (type 2 intrinsic stacking fault of HCP, I_2 (Benoit et al. 2013)), which comprises two layers of atoms surrounded by FCC structures (green atoms) and a low energy HCP structure. Accordingly, the FCC region occupies a significantly less area fraction compared to the HCP region. In addition, a noticeable spiral pattern on the dislocation line around the nodes is clearly visible in Fig. 25c, d, and the same feature was also observed in Cu-Ni interfaces (Shao and Wang 2016).

The energy of the Ti/TiN interfaces, for both ORs, as a function of the L parameter is given in Fig. 26a. The interface energy (γ) is calculated using the following formula:

$$\gamma = \frac{1}{A} [E_{\text{interf.}}(nTi_c, mTi_m, lN) - nE_{Ti}^c - mE_{Ti}^m - lE_N],$$

where A is the area of the interface in the computational cell, $E_{\text{interf.}}$ is the total energy of the structure containing the interface, n is the number of Ti atoms in TiN layer, m is the number of Ti atoms in the Ti layer, and l is the number of N atoms. E_{Ti}^c and E_{Ti}^m are the cohesive energies of the Ti atoms in TiN (ceramic) and Ti (metal). E_N is the cohesive energy of N. It is apparent from Fig. 26a that the interface energy significantly decreases when the MDN is moved away from the Ti/TiN chemical interface by one monolayer of Ti and gradually increases as the MDN is moved further into Ti. As was discussed above, the dramatic reduction in interface energy is associated with the reduction in the MDN's elastic energy as well as the interface stacking fault energy. The gradual increase in interface energy when $L \geq 1$ is associated with the increase in the number of coherent Ti atomic monolayers situated between the MDN and the Ti/TiN chemical interface.

Fig. 26 The variation in interface free energy (a) and the interface shear strength (b) of the Ti/TiN interface as a function of location of MDN (L)



Another observation from Fig. 26a is that the interface energy for OR 2 is consistently smaller than that for OR 1. This is primarily due to the greater area fraction of low energy structures in the OR 2 interfaces. Next, we discuss this in detail, separately, using the cases of $L = 0$ and $L > 0$. For $L = 0$, the OR 1 interface comprises the higher-energy Overlap_S structure and the intermediate-energy FCC_S structure, while the OR 2 interface comprises higher-energy Overlap_S structure and the low-energy HCP_S structure. This gives rise to the significantly higher energy of the OR 1 (-2.62 J/m²) interface compared to OR 2 (-2.88 J/m²). On the other hand, when $L > 0$, the coherent structures in OR 1 comprise only I₁ stacking fault, while the coherent structures in OR 2 comprise I₂ stacking fault as well as perfect HCP stacking. Energetically, according to the current work and the other studies (Benoit et al. 2013), the stacking fault energy of I₁ is around 1/2 of that of the I₂. In fact, it has been reported, agreeing with the current work, $\gamma_{I_1} = 148.6$ mJ/m² and $\gamma_{I_2} = 259.1$ mJ/m² (Benoit et al. 2013). Therefore, the total stacking fault energies for the interfaces are $E_{I_1} = \gamma_{I_1}A$ and $E_{I_2} = \eta\gamma_{I_2}A$, where A is the area of the unit simulation cell, η is the area fraction of the I₂ stacking fault, and $\eta < 0.5$.

It is apparent that $E_{I_1} > E_{I_2}$. Therefore, the interface energy of OR 1 interfaces is always greater than that of the OR 2 interfaces.

It is also interesting to note the correlation between variation of interface free energy (Fig. 26a) and the variation in the theoretical shear strength of the interface (Fig. 26b), as a function of the location of the MDN (L number). As mentioned before, the intrinsic energetic properties of the coherent interface structure (such as the GSF profile and the work of adhesion) affect the exact characteristics of the MDN (such as core widths of the dislocations and nodes) and, ultimately, the overall mechanical response of the Ti/TiN interface. Therefore, in the following sections, we first discuss the intrinsic properties of the coherent Ti/TiN interfacial structures (section “[Intrinsic Energetic Characteristics of Coherent Ti/TiN Interfacial Structures](#)”), then the detailed shear deformation mechanisms of the interfaces are presented (section “[The Shear Response of the Ti–TiN Interfaces](#)”).

Intrinsic Energetic Characteristics of Coherent Ti/TiN Interfacial Structures

The structural characteristics and, in turn, the mechanical properties of the Ti/TiN interfaces are dictated by the intrinsic properties of the coherent interfacial structures. DFT, although computationally limited to length scales of a few Å in each dimension, has great advantages in probing properties of metal/ceramic interfaces. It explicitly models the valence electrons of atoms, which provide a highly accurate assessment of the fundamental properties of coherent interface structures. As will be discussed in the following sections, it is these properties that govern the structure and mechanical response of “real” interfaces between Ti and TiN layers, i.e., the semi-coherent interfaces. The energy characteristics of the fully coherent Ti/TiN chemical interface as well as the nearby Ti (0001) slip planes obtained using DFT calculations (Zhang et al. 2017) are analyzed and presented below.

Shown in Fig. 27 is the plot of the work of adhesion (W_A) on the Ti/TiN chemical interface ($L = 0$), as well as the various (0001) planes in Ti close to the chemical interface ($L > 0$). Clearly due to the strong ionic bonds formed between the Ti and N atoms, W_A on the chemical interface is very large, ~ 6.11 J/m². As the plane of analysis moves away from the chemical interface, W_A decreases dramatically and appears to saturate at 3.74 J/m² when $L \geq 2$. Theoretically, the work of adhesion is related to the free surface energy by factor of two, i.e., $W_A = 2\gamma$. The work of adhesion on the (0001) plane of bulk Ti, shown as the red dashed line in Fig. 27, is calculated to be 4.02 J/m², which agrees well with the experimental measured (0001) surface energy of 2.1 J/m² and 1.92 J/m² (Benoit et al. 2013). This indicates that the cohesive strength of the (0001) Ti reaches a minimum due to the presence of the Ti/TiN chemical interface, below that for bulk Ti.

A similar trend is observed in the generalized stacking fault energy (GSFE) profiles of the Ti/TiN chemical interface ($L = 0$), as well as the Ti (0001) slip planes near the chemical interface ($L > 0$). Shown in Fig. 28a is the GSFE surfaces for $L = 0$. The GSFE are calculated using the minimum energy structures as the reference. The minimum energy structure when at the chemical interface ($L = 0$) correspond to point 1 in Fig. 28a. The minimum energy interface structure is such

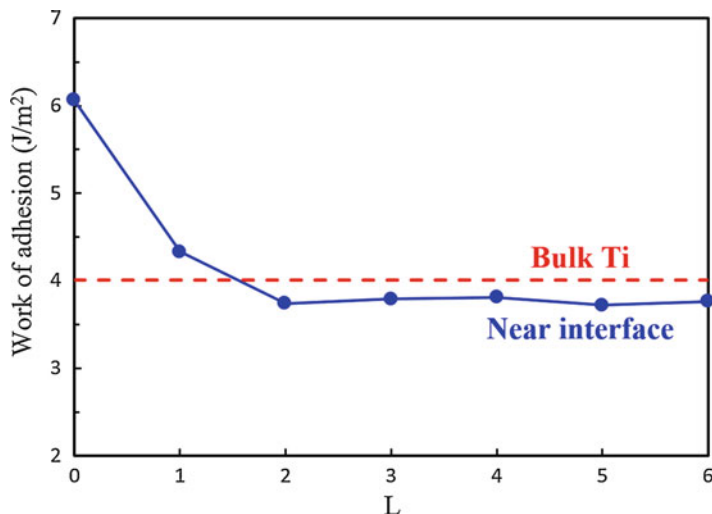


Fig. 27 The variation of the work of adhesion (W_A) on the Ti/TiN chemical interface and various (0001) slip planes of Ti as a function of L number

that the Ti atoms form the stable FCC structure across the interface. The Ti and N atoms immediately adjacent to the interface form the stable NaCl structure. A metastable coherent structure, point 2 on Fig. 28a, also exists on the chemical interface. In this structure, the Ti atoms adjacent to the interface overlap. This metastable structure is named $Overlap_S$. The energy for $Overlap_S$ structure is 0.89 J/m^2 . The maximum of GSFE at the chemical interface (2.71 J/m^2) occurs when, immediately to the interface plane, the Ti atoms are right on top of the N atoms. The maximum GSFE corresponds to point 3 in Fig. 28a and structure 3 in Fig. 28c, where the Ti atoms across the interface form an HCP structure. This structure is unstable and is therefore named HCP_U . Note that maximum stacking fault energy of the chemical interface is so high that it exceeds the (0001) free surface energy of the bulk Ti (2.1 J/m^2) (Benoit et al. 2013).

When $L \geq 1$, the amplitude of the GSFE drastically decreases. For clarity of display, only one typical case ($L = 1$) is shown in Fig. 28b. Note that the energy minimum (Point α in Fig. 28b) corresponds to the stable HCP structure of Ti, which is the reference of the GSFE surface. The metastable structure (I_2 stacking fault of HCP crystals) has GSFE of 0.32 J/m^2 (point β in Fig. 28b). The maximum GSFE for $L = 1$ is 0.63 J/m^2 (Point γ in Fig. 28b). The profile of the GSFE along the $[\bar{1}100]$ direction is shown in Fig. 28d, along with the profile along the same direction on the GSFE surfaces for $L = 2, 3$ and in bulk Ti. The GSFE profile changes slightly from $L = 1$ to $L = 3$. The smallest variation in GSFE occurs when $L = 3$, which is noticeably smaller than in bulk Ti.

The results presented above indicate the existence of a weak interaction plane in the Ti metal near the Ti/TiN chemical interface. As will be discussed in the

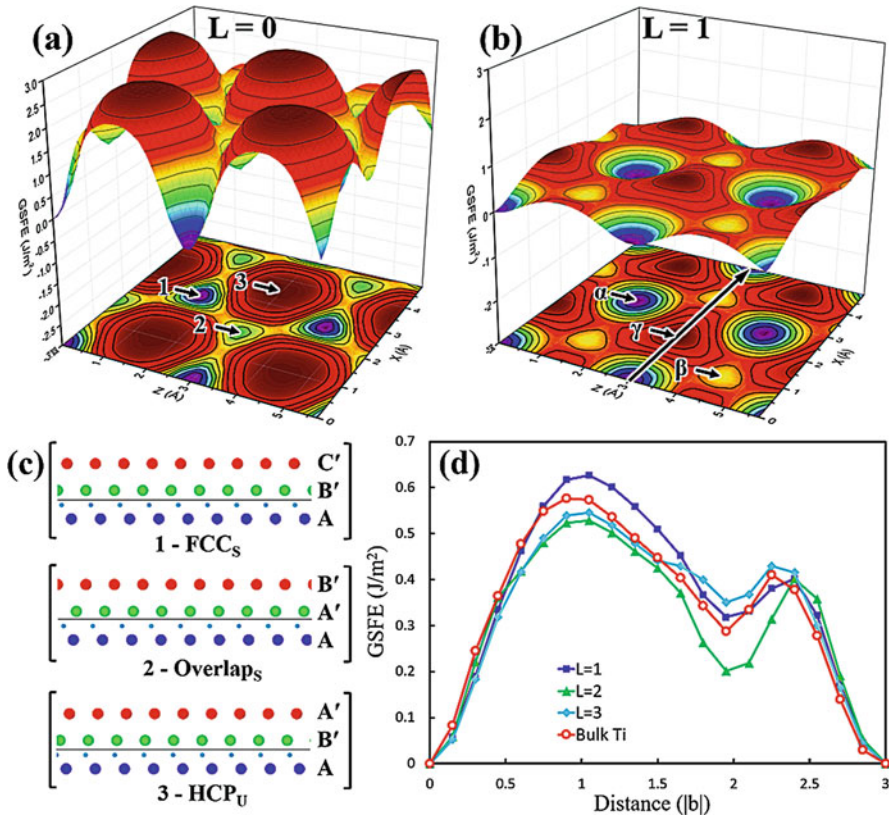


Fig. 28 (a) & (b) The surface plot of the generalized stacking fault energy of the Ti/TiN chemical interface ($L = 0$) as well as the (0001) slip plane of Ti when $L = 1$. (c) The coherent interface structure corresponding to the Points 1, 2, and 3 in (a). (d) The profile of GSFE along $[\bar{1}100]$ for $L = 1, 2, 3$ and bulk Ti

upcoming section, the weakened interaction on that plane alters the structure therefore the mechanical properties of the MDN which negatively impacts the shear strength of the Ti/TiN interface.

The Shear Response of the Ti–TiN Interfaces

As was shown in section “The Location of the Misfit Dislocation Network,” the presence of the chemical interface alters the energy landscape (including the work of adhesion, as well as the generalized stacking fault energy, or the γ surface) of the Ti (0001) slip planes near the Ti/TiN chemical interface. There exists a weak interaction plane inside Ti but near the Ti/TiN chemical interface. Just a small distance away from the chemical interface, the work of adhesion reaches a minimum value of around $3.72 \sim 3.76 \text{ J/m}^2$ when $L \geq 2$, which is significantly less than that in the bulk Ti, which is around 4.02 J/m^2 (Fig. 27). In addition, the amplitude of the

generalized stacking fault energy profile on Ti (0001) plane also reaches a minimum when $L = 3$, which is noticeably smaller than that of bulk Ti (Fig. 28). One expects that the presence of such a “weak” plane would negatively impact the mechanical response of the interface region.

As was discussed in section “The Structure of the Misfit Dislocation Network and the Coherent Regions of the Ti-TiN Interfaces,” the width/diameter of the dislocation/nodal cores is dependent on the amplitude of the GSFE profile. According to GSFE surfaces and profiles obtained from the DFT calculations, the core-size of the dislocations/nodes of the MDN is expected to be the largest when $L = 3$. As was shown by earlier studies, the excessively constricted nodal cores give rise to localization of free volume, which lead to transformation of nodal structures. In fact, the small change in GSFE induced by biaxial tension/compression parallel to the interface is sufficient to produce the concentration of free volume which then lead to transformation of dislocation structure at nodes.

Our MD simulation qualitatively confirmed the predictions above. The dislocation and atomic structures at a node for the cases of $L = 1$ and $L = 5$ are shown in Fig. 29. When $L = 1$ (Fig. 29a), due to the relatively weak interface interaction associated with the GSFE profile with a reduced magnitude, the node is simply an intersection point of dislocations. The size of the node is comparable to the core-width of dislocations. The free volume at the node is smeared, as is evident in the (0001) Ti atomic layer of $M = 2$ shown in Fig. 29c. The atoms are colored according to centro-symmetry analysis performed only on Ti atoms. Blue color indicates a centro-symmetric FCC structure (a stacking fault in Ti); yellow color indicates a nonsymmetric HCP structure; red color indicates a locally disordered structure. When $L > 1$, including in bulk Ti, the increased GSFE profile (as compared to the case of $L = 1$) constricted the size of the node which gives rise to the concentration of the free volume. Fig. 29b, d shows a typical case of $L = 5$. The highly localized free volume is evident in the (0001) atomic plane with $M = 6$ (Fig. 29d). Once again, there is a qualitative agreement between the results from MD and DFT. According to the MD, the (0001) plane that has the smallest amplitude of GSFE is $L = 1$, while DFT predicted $L = 3$. As was discussed by Shao and others (Shao et al. 2015; Wang and Misra 2014; Henager et al. 2004), the concentration of free volume is the result of the presence of three dislocation jogs (marked by three vertical line segments in green color in Fig. 29b) at the node. Such jogs have Burgers vectors $b_4 = \frac{a}{3} [\bar{1}2\bar{1}0]$, $b_5 = \frac{a}{3} [\bar{1}\bar{1}20]$, and $b_6 = \frac{a}{3} [2\bar{1}\bar{1}0]$ and are glissile on the prismatic planes. However, the aggregation of the three jogs serves as strong pinning point to the motion of MDN due to the following two reasons. First, the global loading condition (simple shear parallel to interface) exerts zero resolved shear stress on the prismatic slip systems ($\{10\bar{1}0\} \{1\bar{2}10\}$). To show this, consider the stress applied to the bilayer system:

$$\sigma = \begin{bmatrix} 0 & \tau & 0 \\ \tau & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

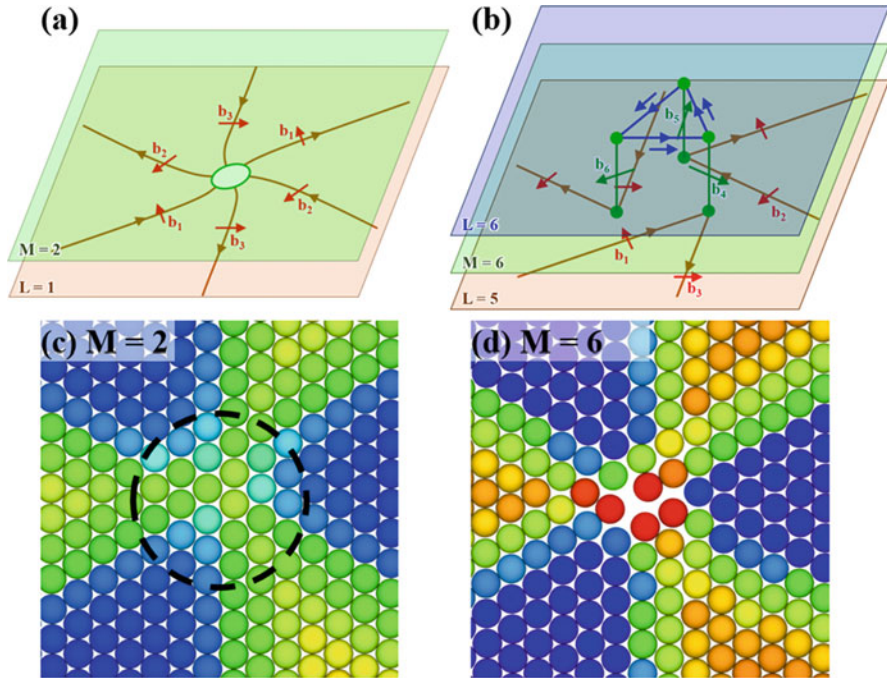


Fig. 29 Dislocation and atomic structures at a node on a Ti (0001) atomic monolayer for $L = 1$ and $L = 5$. (a) and (b) show the dislocation structures with line senses and Burgers vectors indicated. (c) and (d) show the corresponding atomic structures. Note that the atomic monolayers are on the tension side of the MDNs. The atoms are colored according to centro-symmetry analysis performed only on Ti atoms. Blue color indicates a centro-symmetric FCC structure (a stacking fault in Ti); yellow color indicates a non-symmetric HCP structure; red color indicates a locally disordered structure

a random prismatic plane with plane normal:

$$\hat{n} = [n_x \ 0 \ n_z],$$

and an unit vector parallel to $\frac{a}{3}(\bar{1}2\bar{1}0)$ type Burgers vector in the prismatic plane:

$$\hat{v} = [n_z \ 0 \ -n_x].$$

Obviously, since \hat{v} lies within the prismatic plane, $\hat{n} \cdot \hat{v} = 0$. The resolved shear stress of this slip system is $\hat{n} \cdot \sigma \cdot \hat{v} \equiv 0$. Therefore, the jogs do not move in response to the loading. Second, the jogs are connected by Shockley partial dislocation lines (blue dislocation lines forming a triangle in Fig. 29b) residing in the (0001) slip plane ($L = 6$) immediately above the MDN, which also encloses a

stacking fault. As was shown in section “[Mechanistic Understanding on the Shear Instability of Metal/Ceramic Interface Provided by Simulations](#),” the stable stacking fault energy of I_2 structure in Ti is very high, ranging between 0.2 and 0.35 J/m². Driven by the reduction in energy associated with the extra dislocation lines and the stacking fault, the jogs shown in Fig. 29d are tightly bound. Evidently, any motion of the tightly bound jogs leads to increase in the stacking fault energy and dislocation line energy in the blue triangle shown in Fig. 29d. Therefore, the jogs, as an aggregation, are sessile, and its motion is nonconservative and requires climb.

In bulk Ti, the amplitude of the GSFE profile is large enough to induce the formation of a jogged nodal dislocation structure, as predicted by our MD simulations. Qualitatively agreeing with the DFT calculation, the MD simulation predicts that, at a (0001) glide plane close to the chemical interface (e.g., $L = 1$), the GSFE profile is noticeably diminished which transforms the nodes from the jogged structure (Fig. 29b, d) to a smeared structure (Fig. 29a, c). The smeared node structure does not offer pinning effect to the gliding motion of the MDN. The pinning effect of the jogs shown in Fig. 29b, d is evident in Fig. 26b. As is shown, due to the strong interaction at the chemical interface ($L = 0$), the strength of the interface is extremely high. In fact, in this case, interactions at the chemical interface is so strong, that no shear on the $L = 0$ plane is observed. Instead, slip on $L = 1$ plane is observed in the form of nucleation and propagation of dislocations. When the MDN is located at $L > 1$, the pinning effect of the nodes becomes increasingly stronger, which increases the shear strength of the interface. When $L = 6$, the interface shear strength of both ORs approached 2 GPa, approaching that in bulk Ti.

The shear response of the Ti/TiN interface, in terms of the motion/transformation of the MDN, is shown in Fig. 30. As shown, the MDN, which contains edge dislocation lines, moves towards the positive x direction driven by the Peach-Koehler force induced by the applied shear stress. For $L = 1$ (Fig. 30a–c, g–i), the nodes of the MDN offer minimal pinning, and the entire MDN translates approximately in a rigid manner. In response to the loading, the curvature of the dislocation lines seems to increase associated with the more pronounced spiral pattern at nodes. Upon unloading, the spiral pattern does not vanish. This suggests that the more pronounced spiral pattern is energetically more stable, the mechanical loading simply provided additional perturbation to help the system achieve the global energy minimum. When $L > 1$, the formation of the jogs significantly increased the pinning effect of the nodes for both ORs. Shown in Fig. 30d–f, j–l is the shear response of the MDN when $L = 5$. The strong pinning effect of the nodes is reflected by the bowing of partial dislocation lines. Continued loading with increased stress gives rise to “stress assisted climb” of the nodes (Hirth and Lothe 1982), which leaves vacancies on their trace (Fig. 30f, l). This is a further proof that it is indeed the presence of the jogs that is responsible for the pinning effect of the nodes. The pinning effect of the nodes is reflected by the increased shear strength of the interfaces of both ORs as shown in Fig. 26b, i.e., when the

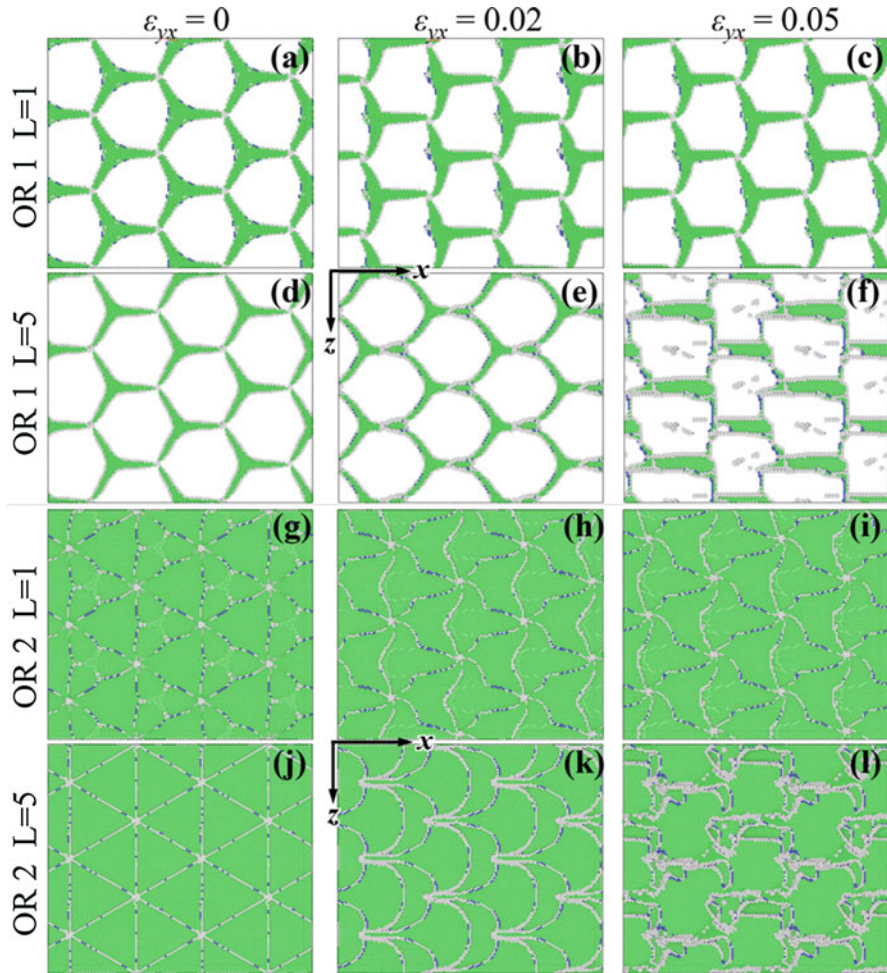


Fig. 30 The sequential snapshots of the MDN at three different effective shear strain levels (0, 0.02 and 0.05) for the two ORs and $L = 1, 5$

location of the MDN changes from $L = 1$ to $L = 6$, the strength increase from 500 to 1000 MPa range to around 2000 MPa. Comparing Fig. 26a, b, it is also worth noting that location of the MDN ($L = 1$) that corresponds to the minimum interface shear strength also corresponds to the minimum interface energy. Our results suggest that the minimum energy Ti/TiN interface has significantly reduced shear strength due to the presence of the MDN. Such metal/ceramic interfaces may exhibit weaker mechanical response (e.g., maximum shear stress) compared to the inside of the metal layer.

Summary

Mechanical integrity of metal/ceramic interfaces is critical for wide ranging applications. The application of ceramic hard coatings on mechanical components, including gears and bearings, can increase their service life by several folds (Jiang et al. 2003). Despite the industrial significance of engineering surfaces of mechanical components and manufacturing tools (Holmberg and Matthews 2009; Mercer et al. 2003; Kotzalas and Doll 2010), engineering of MCIRs still proceeds largely in a trial-and-error manner at the present time. This results in slow development, high cost, inconsistent outcomes, and sometimes failure to deploy. The challenge lies in the incomplete physical understanding and an under-developed defect theory for the mechanical response of MCIRs. The necessary understanding is hampered by (1) a lack of quantitative experimental measurements to assess the mechanical response of the MCIRs; (2) a lack of synergetic, physics-based modeling and simulation efforts to investigate the deformation physics of the MCIRs; (3) a lack of physics-based, experimentally validated, efficient, micro- to mesoscale predictive tools to enable the accelerated discovery-design-implementation cycle for metal/ceramic systems.

To tackle the challenge, research scheme outlined in this chapter centers on understanding the interface failure mechanisms and improving the strength of MCIRs through architectural and chemical refinements following a multiscale ICME approach. The mechanical response and deformation physics of MCIRs with various types of modifications can be investigated by combining first-principles DFT and MD calculations, suggesting experimental designs for strong MCIRs. State-of-the-art vapor phase synthesis in combination with quantitative microscale shear, compression, and tension testing and multiscale materials characterization can be employed to validate model outputs and test MCIR designs. Micro- and mesoscale materials models can then be developed to foster true physics-based MCIR design and accelerate the implementation of next generation applications, e.g., ceramics hard coatings in various surface engineering applications.

In this chapter, we provided an overview on the research on mechanical instability of metal/ceramic interfaces. Challenges and opportunities in the improvements of mechanical integrity of the interfaces have been identified. Theoretical background of this research thrust has been laid out. We concluded by illustrating some preliminary results by the authors on shear instability of various metal/ceramic interfacial regions. In order to achieve the goal of mechanical stability improvement and implementation of next generation MCIRs, true materials-based design, discovery, and understanding of the microscopic deformation physics of MCIRs is crucial. The research scheme described in this chapter outlines an approach to such a problem: (1) developing a multiscale computationally guided, experimentally validated ICME framework that is both time- and cost- efficient; (2) providing fundamental, physics-based understanding that will substantially advance the presently missing defect theories for MCIRs; (3) delivering accurate, experimentally validated, multiscale predictive tools for accelerated MCIR design and implementation.

References

- F. Akasheh, H.M. Zbib, J.P. Hirth, R.G. Hoagland, A. Misra, Dislocation dynamics analysis of dislocation intersections in nanoscale metallic multilayered composites. *J. Appl. Phys.* **101**, 084314 (2007). <https://doi.org/10.1063/1.2721093>
- ASTM International, *ASTM D2095 Standard Test Method for Tensile Strength of Adhesives by Means of Bar and Rod, Annu. B. ASTM Stand.* **15** (2002), pp. 1–3. <https://doi.org/10.1520/D2095-96R15.2>
- ASTM International, *ASTM D1002-05: Standard Test Method for Apparent Shear Strength of Single-Lap-Joint Adhesively Bonded Metal Specimens by Tension Loading (Metal-To-Metal), Standards* (2005), pp. 1–5. <https://doi.org/10.1520/D1002-10.on>
- ASTM International, *ASTM D3165-07 Standard Test Method for Strength Properties of Adhesives in Shear by Tension Loading of Single-Lap-Joint Laminated Assemblies* (2007), p. 4. <https://doi.org/10.1520/D3165-07>
- ASTM International, *ASTM D897 Standard Test Method for Tensile Properties of Adhesive Bonds* (ASTM International, West Conshohocken, 2012), pp. 24–26. <https://doi.org/10.1520/D0897-08R16>
- ASTM International, *Standard Test Method for Strength Properties of Double Lap Shear Adhesive Joints* by (ASTM International, 2013). <https://doi.org/10.1520/D3528-96R08.2>
- ASTM International, *ASTM D4541-17: Standard Test Method for Pull-off Strength of Coatings Using Portable Adhesion Testers* (ASTM international, 2014), pp. 1–16. <https://doi.org/10.1520/D4541-09E01>
- ASTM International, *ASTM C1624-05: Standard Test Method for Adhesion Strength and Mechanical Failure Modes of Ceramic Coatings by Quantitative Single Point Scratch Testing* (ASTM international, 2015)
- M. Benoit, N. Tarrat, J. Morillo, Density functional theory investigations of titanium γ -surfaces and stacking faults. *Model. Simul. Mater. Sci. Eng.* **21**, 015009 (2013). <https://doi.org/10.1088/0965-0393/21/1/015009>
- I.J. Beyerlein, J. Wang, R. Zhang, Interface-dependent nucleation in nanostructured layered composites. *APL Mater.* **1** (2013a). <https://doi.org/10.1063/1.4820424>
- I.J. Beyerlein, J. Wang, R. Zhang, Mapping dislocation nucleation behavior from bimetal interfaces. *Acta Mater.* **61**, 7488–7499 (2013b). <https://doi.org/10.1016/j.actamat.2013.08.061>
- D. Bhattacharyya, N.A.A. Mara, P. Dickerson, R.G.G. Hoagland, A. Misra, Compressive flow behavior of Al–TiN multilayers at nanometer scale layer thickness. *Acta Mater.* **59**, 3804–3816 (2011). <https://doi.org/10.1016/j.actamat.2011.02.036>
- S.D. Brown, Adherence failure and measurement: Some troubling questions. *J. Adhes. Sci. Technol.* **8**, 687–711 (1994). <https://doi.org/10.1163/156856194X00438>
- S.D. Brown, Adherence failure and measurement: Some troubling questions, in *Adhesion Measurement of Films and Coatings*, ed. by K. L. Mittal, 1st edn. (VSP, Utrecht, 1995), pp. 15–39
- S.J. Bull, Can the scratch adhesion test ever be quantitative? *Adhes. Meas. Film. Coatings* **2**, 107 (2001)
- O. Casals, S. Forest, Finite element crystal plasticity analysis of spherical indentation in bulk single crystals and coatings. *Comput. Mater. Sci.* **45**, 774–782 (2009). <https://doi.org/10.1016/j.commatsci.2008.09.030>
- H.Y. Chen, C.J. Tsai, F.H. Lu, The Young's modulus of chromium nitride films. *Surf. Coatings Technol.* **184**, 69–73 (2004). <https://doi.org/10.1016/j.surfcoat.2003.10.064>
- K. Chen, W.J. Meng, J.A. Eastman, Interface development in cu-based structures transient liquid phase (TLP) bonded with thin Al foil intermediate layers. *Metall. Mater. Trans. A.* **45**, 3892–3906 (2014a). <https://doi.org/10.1007/s11661-014-2339-5>
- K. Chen, Y. Mu, W.J. Meng, A new experimental approach for evaluating the mechanical integrity of interfaces between hard coatings and substrates. *MRS Commun.* **4**, 19–23 (2014b). <https://doi.org/10.1557/mrc.2014.3>

- Y. Chen, S. Shao, X.-Y. Liu, S.K. Yadav, N. Li, N. Mara, J. Wang, Misfit dislocation patterns of mg-Nb interfaces. *Acta Mater.* **126**, 552–563 (2017). <https://doi.org/10.1016/j.actamat.2016.12.041>
- B.D. Cullity, S.R. Stock, *Elements of X-Ray Diffraction*, 3rd edn. (Prentice-Hall, New York, 2001)
- R. Darolia, Thermal barrier coatings technology: critical review, progress update, remaining challenges and prospects. *Int. Mater. Rev.* **58**, 315–348 (2013)
- J.A. El-Awady, S. Bulent Biner, N.M. Ghoniem, A self-consistent boundary element, parametric dislocation dynamics formulation of plastic flow in finite volumes. *J. Mech. Phys. Solids* **56**, 2019–2035 (2008). <https://doi.org/10.1016/j.jmps.2007.11.002>
- A.G. Evans, D.R. Clarke, C.G. Levi, The influence of oxides on the performance of advanced gas turbines. *J. Eur. Ceram. Soc.* **28**, 1405–1419 (2008). <https://doi.org/10.1016/j.jeurceramsoc.2007.12.023>
- N.M. Ghoniem, X. Han, Dislocation motion in anisotropic multilayer materials. *Philos. Mag.* **85**, 2809–2830 (2005). <https://doi.org/10.1080/14786430500155338>
- N. Ghoniem, S.-H. Tong, L. Sun, Parametric dislocation dynamics: a thermodynamics-based approach to investigations of mesoscopic plastic deformation. *Phys. Rev. B* **61**, 913–927 (2000). <https://doi.org/10.1103/PhysRevB.61.913>
- N.M. Ghoniem, J. Huang, Z. Wang, Affine covariant-contravariant vector forms for the elastic field of parametric dislocations in isotropic crystals. *Philos. Mag. Lett.* **82**, 55–63 (2002). <https://doi.org/10.1080/09500830110103216>
- S. Groh, E.B. Marin, M.F. Horstemeyer, H.M. Zbib, Multiscale modeling of the plasticity in an aluminum single crystal. *Int. J. Plast.* **25**, 1456–1473 (2009). <https://doi.org/10.1016/j.ijplas.2008.11.003>
- V. Gupta, A.S. Argon, D.M. Parks, J.A. Cornie, Measurement of interface strength by a laser spallation technique. *J. Mech. Phys. Solids* **40**, 141–180 (1992). [https://doi.org/10.1016/0022-5096\(92\)90296-E](https://doi.org/10.1016/0022-5096(92)90296-E)
- V. Gupta, J. Yuan, A. Pronin, Recent developments in the laser spallation technique to measure the interface strength and its relationship to interface toughness with applications to metal/ceramic, ceramic/ceramic and ceramic/polymer interfaces. *J. Adhes. Sci. Technol.* **8**, 713–747 (1994). <https://doi.org/10.1163/156856194X00447>
- V. Gupta, V. Kireev, J. Tian, H. Yoshida, H. Akahoshi, Glass-modified stress waves for adhesion measurement of ultra thin films for device applications. *J. Mech. Phys. Solids* **51**, 1395–1412 (2003). [https://doi.org/10.1016/S0022-5096\(03\)00057-7](https://doi.org/10.1016/S0022-5096(03)00057-7)
- X. Han, N.M. Ghoniem, Stress field and interaction forces of dislocations in anisotropic multilayer thin films. *Philos. Mag.* **85**, 1205–1225 (2005). <https://doi.org/10.1080/14786430412331331907>
- C.H. Henager, R.J. Kurtz, R.G. Hoagland, Interactions of dislocations with disconnections in fcc metallic nanolayered materials. *Philos. Mag.* **84**, 2277–2303 (2004). <https://doi.org/10.1080/14786430410001678235>
- J.P. Hirth, J. Lothe, *Theory of Dislocations*, 2nd edn. (Krieger Publishing Company, Malabar, 1982)
- J.P. Hirth, R. Pond, Steps, dislocations and disconnections as interface defects relating to structure and phase transformations. *Acta Mater.* **44**, 4749–4763 (1996). <http://www.sciencedirect.com/science/article/pii/S1359645496001322>. Accessed 29 Nov 2012
- J.P.P. Hirth, R.C.C. Pond, J. Lothe, Disconnections in tilt walls. *Acta Mater.* **54**, 4237–4245 (2006). <https://doi.org/10.1016/j.actamat.2006.05.017>
- H. Holleck, Material selection for hard coatings. *J. Vac. Sci. Technol. A Vac. Surf. Film* **4**, 2661–2669 (1986). <https://doi.org/10.1116/1.573700>
- K. Holmberg, A. Matthews, *Coatings Tribology – Properties, Mechanisms, Techniques and Applications in Surface Engineering*, 2nd edn. (Elsevier Science, Amsterdam, 2009)
- M.A. Hopcroft, W.D. Nix, T.W. Kenny, What is the Young’s modulus of silicon? *J. Microelectromech. Syst.* **19**, 229–238 (2010). <https://doi.org/10.1109/JMEMS.2009.2039697>
- T. Itoh (ed.), *Ion Beam Assisted Film Growth* (Elsevier, Amsterdam, 1989)

- J.C. Jiang, W.J. Meng, A.G. Evans, C.V. Cooper, Structure and mechanics of W-DLC coated spur gears. *Surf. Coatings Technol.* **176**, 50–56 (2003). [https://doi.org/10.1016/S0257-8972\(03\)00445-6](https://doi.org/10.1016/S0257-8972(03)00445-6)
- Y.M. Kim, B.J. Lee, Modified embedded-atom method interatomic potentials for the Ti-C and Ti-N binary systems. *Acta Mater.* **56**, 3481–3489 (2008a). <https://doi.org/10.1016/j.actamat.2008.03.027>
- Y.-M. Kim, B.-J. Lee, Modified embedded-atom method interatomic potentials for the Ti-C and Ti-N binary systems. *Acta Mater.* **56**, 3481–3489 (2008b). <https://doi.org/10.1016/j.actamat.2008.03.027>
- C. Kittel, Introduction to solid state physics. *Solid State Phys.*, 703 (2005). <https://doi.org/10.1119/1.1974177>
- M.N. Kotzalas, G.L. Doll, Tribological advancements for reliable wind turbine performance. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **368**, 4829–4850 (2010). <https://doi.org/10.1098/rsta.2010.0194>
- C. Kral, W. Lengauer, D. Rafaja, P. Ettmayer, Critical review on the elastic properties of transition metal carbides, nitrides and carbonitrides. *J. Alloys Compd.* **265**, 215–233 (1998). [https://doi.org/10.1016/S0925-8388\(97\)00297-1](https://doi.org/10.1016/S0925-8388(97)00297-1)
- P. Kutilek, J. Miksovsky, The procedure of evaluating the practical adhesion strength of new biocompatible nano- and micro-thin films in accordance with international standards. *Acta Bioeng. Biomech.* **13**, 87–94 (2011)
- R.A. Lebensohn, C.N. Tomé, A self-consistent anisotropic approach for the simulation of plastic deformation and texture development of polycrystals: Application to zirconium alloys. *Acta Metall. Mater.* **41**, 2611–2624 (1993). [https://doi.org/10.1016/0956-7151\(93\)90130-K](https://doi.org/10.1016/0956-7151(93)90130-K)
- R.A. Lebensohn, C.N. Tomé, A self-consistent viscoplastic model: Prediction of rolling textures of anisotropic polycrystals. *Mater. Sci. Eng. A* **175**, 71–82 (1994). [https://doi.org/10.1016/0921-5093\(94\)91047-2](https://doi.org/10.1016/0921-5093(94)91047-2)
- N. Li, H. Wang, A. Misra, J. Wang, In situ nanoindentation study of plastic co-deformation in Al-TiN nanocomposites. *Sci. Rep.* **4**, 6633 (2014). <https://doi.org/10.1038/srep06633>
- J. Lin, W.D. Sproul, J.J. Moore, Microstructure and properties of nanostructured thick CrN coatings. *Mater. Lett.* **89**, 55–58 (2012). <https://doi.org/10.1016/j.matlet.2012.08.060>
- Z. Lin, X. Peng, T. Fu, Y. Zhao, C. Feng, C. Huang, Z. Wang, Atomic structures and electronic properties of interfaces between aluminum and carbides/nitrides: A first-principles study. *Phys. E Low-Dimensional Syst. Nanostruct.* **89**, 15–20 (2017). <https://doi.org/10.1016/j.physe.2017.01.025>
- K. Lukaszewicz, A. Kriz, J. Sondor, Structure and adhesion of thin coatings deposited by PVD technology on the X6CrNiMoTi17-12-2 and X40CrMoV5-1 steel substrates. *Arch. Mater. Sci. Eng.* **51**, 40–47 (2011)
- H. Lyu, A. Ruimi, H.M. Zbib, A dislocation-based model for deformation and size effect in multi-phase steels. *Int. J. Plast.* **72**, 44–59 (2015). <https://doi.org/10.1016/j.ijplas.2015.05.005>
- H. Lyu, N. Taheri-Nassaj, H.M. Zbib, A multiscale gradient-dependent plasticity model for size effects. *Philos. Mag.* **96**, 1883–1908 (2016). <https://doi.org/10.1080/14786435.2016.1180437>
- H. Lyu, M. Hamid, A. Ruimi, H.M. Zbib, Stress/strain gradient plasticity model for size effects in heterogeneous nano-microstructures. *Int. J. Plast.* (2017). <https://doi.org/10.1016/j.ijplas.2017.05.009>
- E.B. Marin, *On the Formulation of a Crystal Plasticity Model* (Livermore, 2006). <http://prod.sandia.gov/techlib/access-control.cgi/2006/064170.pdf>
- P.H. Mayrhofer, G. Tischler, C. Mitterer, Microstructure and mechanical/thermal properties of Cr-N coatings deposited by reactive unbalanced magnetron sputtering. *Surf. Coatings Technol.* **142–144**, 78–84 (2001). [https://doi.org/10.1016/S0257-8972\(01\)01090-8](https://doi.org/10.1016/S0257-8972(01)01090-8)
- W.J. Meng, T.J. Curtis, Inductively coupled plasma assisted physical vapor deposition of titanium nitride coatings. *J. Electron. Mater.* **26**, 1297–1302 (1997)
- W.J. Meng, G.L. Eesley, Growth and mechanical anisotropy of TiN thin films. *Thin Solid Films* **271**, 108–116 (1995). [https://doi.org/10.1016/0040-6090\(95\)06875-9](https://doi.org/10.1016/0040-6090(95)06875-9)

- W.J. Meng, T.J. Curtis, L.E. Rehn, P.M. Baldo, Temperature dependence of inductively coupled plasma assisted growth of TiN thin films. *Surf. Coatings Technol.* **120–121**, 206–212 (1999). [https://doi.org/10.1016/S0257-8972\(99\)00457-0](https://doi.org/10.1016/S0257-8972(99)00457-0)
- W.J. Meng, E.I. Meletis, L.E. Rehn, P.M. Baldo, Inductively coupled plasma assisted deposition and mechanical properties of metal-free and Ti-containing hydrocarbon coatings. *J. Appl. Phys.* **87**, 2840–2848 (2000). <https://doi.org/10.1063/1.372266>
- C. Mercer, A.G. Evans, N. Yao, S. Allameh, C.V. Cooper, Material removal on lubricated steel gears with W-DLC-coated surfaces. *Surf. Coatings Technol.* **173**, 122–129 (2003). [https://doi.org/10.1016/S0257-8972\(03\)00467-5](https://doi.org/10.1016/S0257-8972(03)00467-5)
- M.A. Meyers, K.K. Chawla, *Mechanical Behavior of Materials* (Cambridge University Press, 2007)
- R.E. Miller, L. Shilkrot, W.A. Curtin, A coupled atomistics and discrete dislocation plasticity simulation of nanoindentation into single crystal thin films. *Acta Mater.* **52**, 271–284 (2004). <https://doi.org/10.1016/j.actamat.2003.09.011>
- Y. Mu, J.W. Hutchinson, W.J. Meng, Micro-pillar measurements of plasticity in confined cu thin films. *Extrem. Mech. Lett.* **1**, 62–69 (2014). <https://doi.org/10.1016/j.eml.2014.12.001>
- F.R.N. Nabarro, Dislocations in a simple cubic lattice. *Proc. Phys. Soc.* **59**, 256–272 (1947). <https://doi.org/10.1088/0959-5309/59/2/309>
- J.F. Nye, *Physical Properties of Crystals: Their Representation by Tensors and Matrices* (Oxford University Press, New York, 2000)
- W. Oliver, G. Pharr, An improved technique for determining hardness and elastic modulus using load and displacement-sensing indentation systems. *J. Mater. Res.* **7**, 1564–1583 (1992). <https://doi.org/10.1557/JMR.1992.1564>
- R. Peierls, The size of a dislocation. *Proc. Phys. Soc.* **52**, 34–37 (1940). <https://doi.org/10.1088/0959-5309/52/1/305>
- A.J. Perry, J.A. Sue, P.J. Martin, Practical measurement of the residual stress in coatings. *Surf. Coatings Technol.* **81**, 17–28 (1996). [https://doi.org/10.1016/0257-8972\(95\)02531-6](https://doi.org/10.1016/0257-8972(95)02531-6)
- I. Salehinia, S. Shao, J. Wang, H.M.M. Zbib, Plastic deformation of metal/ceramic Nanolayered composites. *JOM* **66**, 2078–2085 (2014). <https://doi.org/10.1007/s11837-014-1132-7>
- I. Salehinia, S. Shao, J. Wang, H.M. Zbib, Interface structure and the inception of plasticity in Nb/NbC nanolayered composites. *Acta Mater.* **86** (2015). <https://doi.org/10.1016/j.actamat.2014.12.026>
- C. Sant, M. Ben Daia, P. Aubert, S. Labdi, P. Houdy, Interface effect on tribological properties of titanium–titanium nitride nanolaminated structures. *Surf. Coatings Technol.* **127**, 167–173 (2000). [https://doi.org/10.1016/S0257-8972\(00\)00663-0](https://doi.org/10.1016/S0257-8972(00)00663-0)
- S. Shao, J. Wang, Relaxation, structure, and properties of semicoherent interfaces. *JOM* **68**, 242–252 (2016). <https://doi.org/10.1007/s11837-015-1691-2>
- S. Shao, J. Wang, A. Misra, R.G. Hoagland, Spiral patterns of dislocations at nodes in (111) semi-coherent FCC interfaces. *Sci. Rep.* **3** (2013). <https://doi.org/10.1038/srep02448>
- S. Shao, J. Wang, A. Misra, Energy minimization mechanisms of semi-coherent interfaces. *J. Appl. Phys.* **116**, 023508 (2014). <https://doi.org/10.1063/1.4889927>
- S. Shao, J. Wang, I.J. Beyerlein, A. Misra, Glide dislocation nucleation from dislocation nodes at semi-coherent {111} cu–Ni interfaces. *Acta Mater.* **98**, 206–220 (2015). <https://doi.org/10.1016/j.actamat.2015.07.044>
- S. Shao, A. Misra, H. Huang, J. Wang, Micro-scale modeling of interface-dominated mechanical behavior. *J. Mater. Sci.* (2017). <https://doi.org/10.1007/s10853-017-1662-9>
- S. Shao, F. Akasheh, J. Wang, Y. Liu, Alternative misfit dislocations pattern in semi-coherent FCC {100} interfaces. *Acta Mater.* **144**, 177–186 (2018). <https://doi.org/10.1016/j.actamat.2017.10.052>
- S.A. Skirlo, M.J. Demkowicz, Viscoelasticity of stepped interfaces. *Appl. Phys. Lett.* **103**, 171908 (2013). <https://doi.org/10.1063/1.4827103>
- J.R. Smith, T. Hong, D.J. Srolovitz, Metal-ceramic adhesion and the Harris functional. *Phys. Rev. Lett.* **72**, 4021–4024 (1994). <https://doi.org/10.1103/PhysRevLett.72.4021>

- M.E. Straumanis, L.S. Yu, Lattice parameters, densities, expansion coefficients and perfection of structure of Cu and of Cu-in α phase. *Acta Crystallogr. Sect. A*. **25**, 676–682 (1969). <https://doi.org/10.1107/S0567739469001549>
- J.A. Sue, A.J. Perry, J. Vetter, Young's modulus and stress of CrN deposited by cathodic vacuum arc evaporation. *Surf. Coatings Technol.* **68–69**, 126–130 (1994). [https://doi.org/10.1016/0257-8972\(94\)90149-X](https://doi.org/10.1016/0257-8972(94)90149-X)
- T. Sun, X. Wu, R. Wang, W. Li, Q. Liu, First-principles study on the adhesive properties of Al/TiC interfaces: Revisited. *Comput. Mater. Sci.* **126**, 108–120 (2017). <https://doi.org/10.1016/j.commatsci.2016.09.024>
- K.H. Thulasi Raman, M.S.R.N. Kiran, U. Ramamurty, G. Mohan Rao, Structural and mechanical properties of room temperature sputter deposited CrN coatings. *Mater. Res. Bull.* **47**, 4463–4466 (2012). <https://doi.org/10.1016/j.materresbull.2012.09.051>
- L. Toth, *Transition Metal Carbides and Nitrides* (Elsevier Science, 1971)
- M.D. Uchic, D.M. Dimiduk, J.N. Florando, W.D. Nix, Sample dimensions influence strength and crystal plasticity. *Science* **305**, 986–989 (2004). <https://doi.org/10.1126/science.1098993>
- M.D. Uchic, P.A. Shade, D.M. Dimiduk, Plasticity of micrometer-scale single crystals in compression. *Annu. Rev. Mater. Res.* **39**, 361–386 (2009). <https://doi.org/10.1146/annurev-matsci-082908-145422>
- L.J. Vandeperre, F. Giuliani, S.J. Lloyd, W.J. Clegg, The hardness of silicon and germanium. *Acta Mater.* **55**, 6307–6315 (2007). <https://doi.org/10.1016/j.actamat.2007.07.036>
- C.A. Volkert, A.M. Minor, Focused ion beam microscopy and micromachining. *MRS Bull.* **32**, 389–399 (2007). <https://doi.org/10.1557/mrs2007.62>
- J. L. Vossen, W. Kern (eds.), *Thin Film Processes II* (Academic Press, Boston, 1991)
- J. Wang, A. Misra, An overview of interface-dominated deformation mechanisms in metallic multilayers. *Curr. Opin. Solid State Mater. Sci.* **15**, 20–28 (2011). <https://doi.org/10.1016/j.cossms.2010.09.002>
- J. Wang, A. Misra, Strain hardening in nanolayered thin films. *Curr. Opin. Solid State Mater. Sci.* **18**, 19–28 (2014). <https://doi.org/10.1016/j.cossms.2013.10.003>
- J. Wang, R.G. Hoagland, J.P. Hirth, A. Misra, Atomistic simulations of the shear strength and sliding mechanisms of copper–niobium interfaces. *Acta Mater.* **56**, 3109–3119 (2008). <https://doi.org/10.1016/j.actamat.2008.03.003>
- H. Wang, P.D. Wu, C.N. Tomé, Y. Huang, A finite strain elastic-viscoplastic self-consistent model for polycrystalline materials. *J. Mech. Phys. Solids* **58**, 594–612 (2010). <https://doi.org/10.1016/j.jmps.2010.01.004>
- J. Wang, C. Zhou, I.J.I.J. Beyerlein, S. Shao, Modeling interface-dominated mechanical behavior of nanolayered crystalline composites. *JOM* **66**, 102–113 (2014). <https://doi.org/10.1007/s11837-013-0808-8>
- J. Wang, Q. Zhou, S. Shao, A. Misra, Strength and plasticity of nanolaminated materials. *Mater. Res. Lett.* **5** (2017). <https://doi.org/10.1080/21663831.2016.1225321>
- S.K.K. Yadav, S. Shao, J. Wang, X.-Y.X.-Y. Liu, Structural modifications due to interface chemistry at metal-nitride interfaces. *Sci. Rep.* **5**, 17380 (2015). <https://doi.org/10.1038/srep17380>
- W. Yang, G. Ayoub, I. Salehinia, B. Mansoor, H. Zbib, Deformation mechanisms in Ti/TiN multilayer under compressive loading. *Acta Mater.* **122**, 99–108 (2017). <https://doi.org/10.1016/j.actamat.2016.09.039>
- H.M. Zbib, T. Diaz de la Rubia, A multiscale model of plasticity. *Int. J. Plast.* **18**, 1133–1163 (2002). [https://doi.org/10.1016/S0749-6419\(01\)00044-4](https://doi.org/10.1016/S0749-6419(01)00044-4)
- H.M. Zbib, M. Rhee, J.P. Hirth, On plastic deformation and the dynamics of 3D dislocations. *Int. J. Mech. Sci.* **40**, 113–127 (1998). [https://doi.org/10.1016/S0020-7403\(97\)00043-X](https://doi.org/10.1016/S0020-7403(97)00043-X)
- J.M. Zhang, Y. Zhang, K.W. Xu, V. Ji, Representation surfaces of Young's modulus and Poisson's ratio for BCC transition metals. *Phys. B Condens. Matter* **390**, 106–111 (2007). <https://doi.org/10.1016/j.physb.2006.08.008>
- Z.G. Zhang, O. Rapaud, N. Bonasso, D. Mercs, C. Dong, C. Coddet, Control of microstructures and properties of dc magnetron sputtering deposited chromium nitride films. *Vacuum* **82**, 501–509 (2008). <https://doi.org/10.1016/j.vacuum.2007.08.009>

- R.F. Zhang, T.C. Germann, J. Wang, X.-Y. Liu, I.J. Beyerlein, Role of interface structure on the plastic response of cu/Nb nanolaminates under shock compression: Non-equilibrium molecular dynamics simulations. *Scr. Mater.* **68**, 114–117 (2013). <https://doi.org/10.1016/j.scriptamat.2012.09.022>
- R.F. Zhang, I.J. Beyerlein, S.J. Zheng, S.H. Zhang, A. Stukowski, T.C. Germann, Manipulating dislocation nucleation and shear resistance of bimetal interfaces by atomic steps. *Acta Mater.* **113**, 194–205 (2016). <https://doi.org/10.1016/j.actamat.2016.05.015>
- X. Zhang, B. Zhang, Y. Mu, S. Shao, C.D.C.D. Wick, B.R. Ramachandran, W.J. Meng, Mechanical failure of metal/ceramic interfacial regions under shear loading. *Acta Mater.* **138**, 224–236 (2017). <https://doi.org/10.1016/j.actamat.2017.07.053>



Uniqueness of Elastoplastic Properties Measured by Instrumented Indentation

6

L. Liu, Xi Chen, N. Ogasawara, and N. Chiba

Contents

Introduction	212
Challenging the Uniqueness of Indentation Load-Displacement Curve vs. Material Property	214
Computation Method	214
Determine Special Materials with Same Loading Curves for Dual Sharp Indenters	215
A Simple Relationship Between the Loading Curvature and Material Properties	215
Special Materials with Same Loading Curvature for Dual Sharp Indenters	216
Determine Special Materials with Same Loading and Unloading Curves for one Conical Indenter	218
Relating the Unloading Work with Material Properties	218
Special Materials with Same Loading and Unloading Curves for a Sharp Indenter	220
Determine Mystical Materials with Same Loading and Unloading Curve for Dual Indenters	222
Weak-Form Mystical Materials and Their Possible Existence	222
Search for Mystical Materials with Fixed Poisson's Ratio	225

L. Liu (✉)

Department of Mechanical and Aerospace Engineering, Utah State University, Logan, UT, USA
e-mail: ling.liu@usu.edu

X. Chen (✉)

Department of Earth and Environmental Engineering, Columbia Nanomechanics Research Center, Columbia University, New York, NY, USA
e-mail: xichen@columbia.edu

N. Ogasawara (✉)

Department of Mechanical Engineering, National Defense Academy of Japan, Yokosuka, Japan
e-mail: oga@nda.ac.jp

N. Chiba

National Defense Academy of Japan, Yokosuka, Japan
e-mail: chiba@nda.ac.jp

Alternative Methods to Distinguish Mystical Materials	231
Improved Spherical Indentation	231
Film Indentation	232
Detectable Strain Range of Indentation Test	234
The Critical Strain	234
Variation of Critical Strain: A Qualitative Explanation	237
Conclusion	238
References	239

Abstract

Indentation is widely used to extract material elastoplastic properties from the measured load-displacement curves. One of the most well-established indentation technique utilizes dual (or plural) sharp indenters (which have different apex angles) to deduce key parameters such as the elastic modulus, yield stress, and work-hardening exponent for materials that obey the power-law constitutive relationship. Here we show the existence of “mystical materials,” which have distinct elastoplastic properties, yet they yield almost identical indentation behaviors, even when the indenter angle is varied in a large range. These mystical materials are, therefore, indistinguishable by many existing indentation analyses unless extreme (and often impractical) indenter angles are used. Explicit procedures of deriving these mystical materials are established, and the general characteristics of the mystical materials are discussed. In many cases, for a given indenter angle range, a material would have infinite numbers of mystical siblings, and the existence maps of the mystical materials are also obtained. Furthermore, we propose two alternative techniques to effectively distinguish these mystical materials. In addition, a critical strain is identified as the upper bound of the detectable range of indentation, and moderate tailoring of the constitutive behavior beyond this range cannot be effectively detected by the reverse analysis of the load-displacement curve. The topics in this chapter address the important question of the uniqueness of indentation test, as well as providing useful guidelines to properly use the indentation technique to measure material elastoplastic properties.

Keywords

Indentation · Elastoplastic properties · Unique solution · Numerical study · Indistinguishable load-displacement curve · Reverse analysis · Detectable strain range · Critical strain · Loading curvature · Indenter angle

Introduction

Instrumented indentation is widely used to probe the constitutive relationships of engineering materials. Without losing generality, the uniaxial true stress-strain curve of a stress-free elastoplastic solid can be expressed in a power-law form, which is a

good approximation for most metals and alloys (Cheng and Cheng 2004).

$$\sigma = E\varepsilon \text{ for } \varepsilon \leq \frac{\sigma_y}{E} \text{ and } \sigma = \sigma_y \left(\frac{E}{\sigma_y} \right)^n \varepsilon^n = R\varepsilon^n, \quad (1)$$

where E is the Young's modulus, σ_y is the yield stress, and n is the work-hardening exponent. For most metals and alloys, n is between 0.0 and 0.5, E is between 10 and 600 GPa, σ_y is between 10 and 2000 MPa, and E/σ_y is between 100 and 5000 (Ashby 1999) – this is a technical range of the engineering materials suitable for the conventional indentation analysis where finite strains are involved.

Four independent parameters (E , ν , σ_y , n) are needed to completely characterize the elastoplastic properties of a power-law stress-free material. Probing these material parameters by indentation has become a focal point of interest in the indentation literature, and various techniques were proposed; see the review by Cheng and Cheng (2004). However, even for some of the existing techniques that are considered as “well-established,” the fundamental question of whether the elastoplastic properties of a specimen can be uniquely determined is still open.

For any indentation technique, the existence of a unique solution requires that the indentation response must be unique for a given material, i.e., one-to-one correspondence between the shape factors of the measured indentation load-displacement curves and material elastoplastic properties. For example, when the apex angle of a sharp indenter is fixed, several research groups have shown that a set of special materials with distinct elastoplastic properties may yield almost the same indentation load-displacement curves (Cheng and Cheng 1999; Capehart and Cheng 2003; Tho et al. 2004; Alkorta et al. 2005). Therefore, the mechanical properties of these specimens cannot be uniquely determined by using one sharp indenter.

In this chapter based on Chen et al. (2007) and Liu et al. (2009), we carry out a systematic numerical study to correlate the indentation responses with a wide range of material properties and a variety of indenter geometries and present an explicit formulation to determine the special sets of materials with distinct elastoplastic properties yet exhibit indistinguishable indentation behaviors even when different indenters are employed. We call such sets of special materials as *mystical materials* – i.e., they are beyond the previous (and conventional) understanding of this topic. For many power-law materials, they have infinite mystical siblings that have indistinguishable loading and unloading behaviors during the indentation test, for a wide range of indenter geometries. Due to the lack of unique solutions, theoretically, these mystical materials are unable to be distinguished by many previously established techniques, including the dual (or plural) sharp indentation method and the conventional spherical indentation analysis with small penetration. The existence map of and the common characteristics of the mystical materials are also established. We then illustrate that the properties of these mystical materials may still be distinguishable by alternative indentation techniques, for example, a film indentation analysis (Zhao et al. 2006a) and an improved spherical indentation technique (Zhao et al. 2006a), which are suitable for specimens with finite thickness and bulk materials, respectively.

Challenging the Uniqueness of Indentation Load-Displacement Curve vs. Material Property

Most indentation techniques, the sharp indentation analysis, the spherical indentation analysis, the film indentation analysis, etc., in essence, require numerical analyses to correlate various shape factors of the P - δ curve with the specimen elastoplastic properties. In doing so, a critical theoretical question emerges: is there a one-to-one correspondence between indentation load-displacement curves (for loading and/or unloading) and material properties (E , ν , σ_y , n)? The uniqueness of the solution of the reverse analysis is the key verification for all indentation analyses – although some of those methods are now widely used in practice and cited in literature, unfortunately, the uniqueness of their solutions has been rarely challenged.

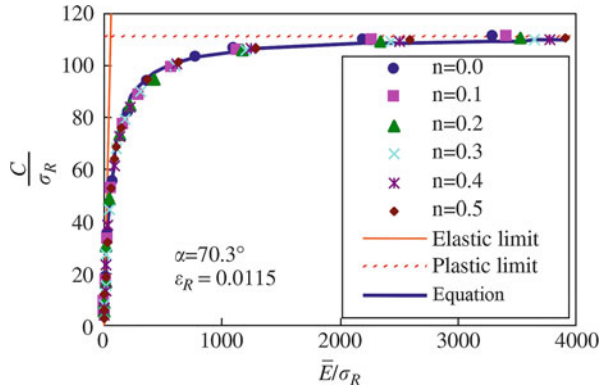
The fundamental question is, does a set of mystical materials which will yield indistinguishable indentation load-displacement curves for not only one particular indenter angle, but also another indenter angle exist? If this is true, then the dual (or plural) indenter method, regardless of the detail of the theory, cannot be used to distinguish these mystical materials. Such a fundamental question is not only the basis for the dual (or plural) sharp indenter method but also the foundation of the spherical indentation and film indentation techniques, as well as most indentation analyses which rely on the load-displacement curves.

In what follows, we will first present a simple and explicit technique to derive special sets of materials (with different elastoplastic properties) that lead to the same loading curves during sharp indentation when different indenter angles are used. Next, we show another technique to derive special sets of materials that yield almost same loading and unloading curves for a given indenter angle. We then extend our analysis to predict the mystical materials that have indistinguishable loading and unloading curves when different sharp indenters are used. Consequently, many of the previously established indentation analyses would fail to distinguish these materials. The existence range, trend, and special features of the mystical materials are discussed. We also show that these mystical materials may still be distinguished by using the improved spherical indentation and film indentation techniques.

Computation Method

In this chapter, the relationships between indentation responses, material properties, and indenter geometries are established from extensive finite element analyses. FEM calculations are performed using the commercial code. The rigid contact surface option is used to simulate the rigid indenter, and the option for finite deformation and strain is employed. A typical mesh for the axisymmetric indentation model comprises about 10,000 4-node elements with reduced integration. The Coulomb's friction law is used between contact surfaces, and the friction coefficient is taken to be 0.15 (Bowden and Tabor 1950), which is a minor factor for indentation (Mesarovic and Fleck 1999; Cheng and Cheng 2004) as long as this value is relatively small. The strain gradient effect is ignored by assuming that the indentation

Fig. 1 The relationship between C/σ_R and \bar{E}/σ_R as n is varied, for $\alpha = 70.3$. The numerical results (symbols) are shown with both elastic and rigid plastic limits, and the empirical fitting function incorporating these limits is given in Eq. 6 (which can be extended to other angles) (Chen et al. 2007)



depth is sufficiently deep so that the continuum mechanics still applies to the bulk specimen. In addition, the strain rate effect is also ignored. In order to obtain both complete and robust numerical results, the material parameters are varied over a large range to cover essentially all engineering materials with $\bar{E}/\sigma_R = 3 - 3900$ and $n = 0-0.5$, where \bar{E} is the plane-strain modulus and σ_R is the representative stress, see (Ogasawara et al. 2005, 2007a) for details. And, for the same reason, a large indenter angle range is also used, from 60° to 80° , which covers most of the angle range used in literature. To be consistent with most literature, Poisson’s ratio is fixed at 0.3.

Determine Special Materials with Same Loading Curves for Dual Sharp Indenters

A Simple Relationship Between the Loading Curvature and Material Properties

When a sharp indenter is penetrating a bulk specimen, the loading $P-\delta$ relationship is always quadratic, i.e., $P=C\delta^2$. Therefore, two materials must have the same loading curvature C in order to have the same loading $P-\delta$ curve. A simple relationship, not a high-order fitting polynomial function, between C and \bar{E}/σ_R must be established on a physical basis.

Note that the specimen is essentially elastic when the variable $\bar{E}/\sigma_R \Rightarrow 0$, whereas the material approaches to rigid plastic when $\bar{E}/\sigma_R \Rightarrow \infty$ – the functional form $\prod(\bar{E}/\sigma_R)$ should incorporate the limits of both mechanisms, such that it remains valid for all materials regardless of the range of data used for fitting. In the representative case of $\alpha = 70.3$ (Fig. 1), both of the elastic and rigid plastic limits can be well defined:

(a) When $\bar{E}/\sigma_R \Rightarrow 0$ (elastic), $C/\sigma_R = \prod(\bar{E}/\sigma_R)$ varies linearly with \bar{E}/σ_R :

$$\frac{C}{\sigma_R} = m_e \frac{\bar{E}}{\sigma_R} \text{ as } \bar{E}/\sigma_R \Rightarrow 0, \tag{2}$$

where m_e can be derived from the classic solution of indentation on elastic materials (Sneddon 1965), and it agrees well with FEM calculations (Ogasawara et al. 2006):

$$m_e = \frac{2\gamma \tan \alpha}{\pi}. \quad (3)$$

(b) When $\bar{E}/\sigma_R \Rightarrow \infty$ (rigid plastic), C/σ_R approaches a constant:

$$\frac{C}{\sigma_R} = m_p \text{ as } \bar{E}/\sigma_R \Rightarrow \infty, \quad (4)$$

where m_p is the rigid plastic limit of conical indentation into a material that obeys the Mises yield criterion.

$$m_p = 13.2 \tan^2 \alpha + 6.18 \tan \alpha - 8.54 \quad (5)$$

for $50^\circ \leq \alpha \leq 80^\circ$. m_p is equal to 112.1 for the Berkovich indenter. In view of the importance of these two limits, we have proposed a very simple empirical form of $\Pi(\bar{E}/\sigma_R)$ to incorporate both limits (Ogasawara et al. 2006, 2007a):

$$\Pi = \frac{C}{\sigma_R} = \left(\frac{1}{m_e \frac{\bar{E}}{\sigma_R}} + \frac{1}{m_p} \right)^{-1}, \quad (6)$$

from which the representative stress can be obtained as $\sigma_R = m_e C \bar{E} / m_p (m_e \bar{E} - C)$ without iteration. The above equation not only incorporates the elastic and plastic limits (thus having physical meaning and wider range of application), but it also involves no fitting parameter if m_p could be solved analytically.

Special Materials with Same Loading Curvature for Dual Sharp Indenters

With the simple Eq. 6 relating C and material properties, it is now possible to explicitly derive material combinations that have the same loading curvature. First consider the case with one indenter (#A) whose half-apex angle is fixed: once α^A is specified, its related elastic limit m_e^A , rigid plastic limit m_p^A , and representative strain $\varepsilon_R^A = 0.0319 \cot \alpha^A$ can be fixed. For two materials, #1 with elastoplastic property (E_1, σ_{y1}, n_1) and #2 with elastoplastic property (E_2, σ_{y2}, n_2), to have the same loading P - δ curve they must satisfy

$$C_1^A = \frac{m_e^A m_p^A \bar{E}_1}{m_e^A \frac{\bar{E}_1}{\sigma_{R1}^A} + m_p^A} = C_2^A = \frac{m_e^A m_p^A \bar{E}_2}{m_e^A \frac{\bar{E}_2}{\sigma_{R2}^A} + m_p^A}, \quad (7)$$

where the representative stresses are

$$\sigma_{R1}^A = R_1 \left(2 \frac{\sigma_{R1}^A}{E_1} + 2\varepsilon_R^A \right)^{n_1} \quad (8)$$

and

$$\sigma_{R2}^A = R_2 \left(2 \frac{\sigma_{R2}^A}{E_2} + 2\varepsilon_R^A \right)^{n_2} \quad (9)$$

respectively, and they need to satisfy

$$\frac{1}{\sigma_{R2}^A} = \frac{1}{\sigma_{R1}^A} + \frac{m_p^A}{m_e^A} \left(\frac{1}{E_1} - \frac{1}{E_2} \right). \quad (10)$$

Similarly, for another sharp indenter (#B) with a different angle α^B , its elastic limit is m_e^B , rigid plastic limit is m_p^B , and representative strain is ε_R^B . If the two materials will again have the same loading curvature, their representative stresses need to satisfy

$$\frac{1}{\sigma_{R2}^B} = \frac{1}{\sigma_{R1}^B} + \frac{m_p^B}{m_e^B} \left(\frac{1}{E_1} - \frac{1}{E_2} \right) \quad (11)$$

with

$$\sigma_{R1}^B = R_1 \left(2 \frac{\sigma_{R1}^B}{E_1} + 2\varepsilon_R^B \right)^{n_1} \quad (12)$$

and

$$\sigma_{R2}^B = R_2 \left(2 \frac{\sigma_{R2}^B}{E_2} + 2\varepsilon_R^B \right)^{n_2}. \quad (13)$$

The procedure of deriving two materials with different elastoplastic properties yet with the same loading curvature (for both indenters #A and #B) can be concluded as

- (a) Choose any E_1 and E_2 that are different (with fixed $\nu_1 = \nu_2 = 0.3$).
- (b) Choose any value of σ_{y1} and n_1 , and derive $R_1 = \sigma_{y1} (E_1 / \sigma_{y1})^{n_1}$.
- (c) Calculate σ_{R1}^A from Eq. 8 and solve for σ_{R2}^A from Eq. 10.
- (d) Obtain one flow stress-total strain pair of the uniaxial stress-strain curve for material #2 as $\sigma_2^A = \sigma_{R2}^A$ and $\varepsilon_2^A = 2\varepsilon_R^A + 2\sigma_{R2}^A / E_2$.
- (e) Calculate σ_{R1}^B from Eq. 12 and solve for σ_{R2}^B from Eq. 13.
- (f) Obtain another flow stress-total strain pair for material #2 as $\sigma_2^B = \sigma_{R2}^B$ and $\varepsilon_2^B = 2\varepsilon_R^B + 2\sigma_{R2}^B / E_2$.

- (g) From both flow stress-total strain pairs, solve $n_2 = \ln(\sigma_2^A/\sigma_2^B) / \ln(\varepsilon_2^A/\varepsilon_2^B)$, $R_2 = \sigma_2^A/(\varepsilon_2^A)^{n_2}$, and $\sigma_{y2} = \left((E_2)^{n_2/(n_2-1)} / (R_2)^{1/(n_2-1)} \right)$. Finally, from the numerical indentation test, confirm that $C_1^A = C_2^A$ and $C_1^B = C_2^B$.

Therefore, for any given material #1 with elastoplastic property (E_1, σ_{y1}, n_1) , we can explicitly derive a special material #2 with elastoplastic property (E_2, σ_{y2}, n_2) such that they not only yield the same loading curvature when indenter #A is used but also have the same loading curvature when indenter #B is used. There are infinite numbers of such special siblings. The procedure outlined above can be readily extended to identify materials with indistinguishable indentation loading P - δ curves for three different sharp indenters (#A, #B, #C).

An example of the set of special materials is given in Fig. 2: for the five materials (mat1–mat5) that have distinct elastoplastic properties, their uniaxial stress-strain relationships are given in the inset of Fig. 2. These materials not only have the same loading curvature for the Berkovich indenter but are also the same when $\alpha=63.14^\circ$, 75.79° , and 80.0° are used; moreover, for any indenter angle between 63.14° and 80° , their loading curvatures are also the same. From Eqs. 10 and 11, if $\bar{E}_1 > \bar{E}_2$ then $\sigma_{R1}^A < \sigma_{R2}^A$ and $\sigma_{R1}^B < \sigma_{R2}^B$ – this can be verified from the inset where for the special sets of materials, the ones with larger moduli have smaller representative stresses. Moreover, when $\alpha^B > \alpha^A$, we have $\varepsilon_R^B < \varepsilon_R^A$ and $m_p^B/m_e^B < m_p^A/m_e^A$. Therefore, for a pair of such special materials, if $\bar{E}_1 > \bar{E}_2$, the difference between σ_1^B and σ_2^B is larger than that of σ_1^A and σ_2^A , and thus $n_1 > n_2$ and $\sigma_{y1} < \sigma_{y2}$, all can be verified from Fig. 2. In this case, the uniaxial stress-strain curves of these special two materials must intersect outside ε_2^A .

From Fig. 2, during unloading the contact stiffness (and thus unloading work) of these special materials are different, which means that if their Young's moduli are known (with a fixed ν), their plastic properties (σ_y, n) can still be uniquely determined from the loading curvature by using the dual (or plural) indenter method, under the important premise that the two indenter angles are distinct enough such that the two determined total strains ε_2^A and ε_2^B are separated sufficiently apart (to ensure numerical accuracy) (Ogasawara et al. 2006).

Determine Special Materials with Same Loading and Unloading Curves for one Conical Indenter

Relating the Unloading Work with Material Properties

While the normalized C (or equivalently, loading work W_l) is the only shape factor during loading, in principle, there are three shape factors for unloading: the normalized δ_f , contact stiffness S , and unloading work W_e . However, only one of them is completely independent; this is because the curvature of the unloading curve of conical indentation (c.f. Fig. 2) is usually very small (except by the end of the unloading process). Therefore, if either one of the variables (δ_f, S, W_e) is

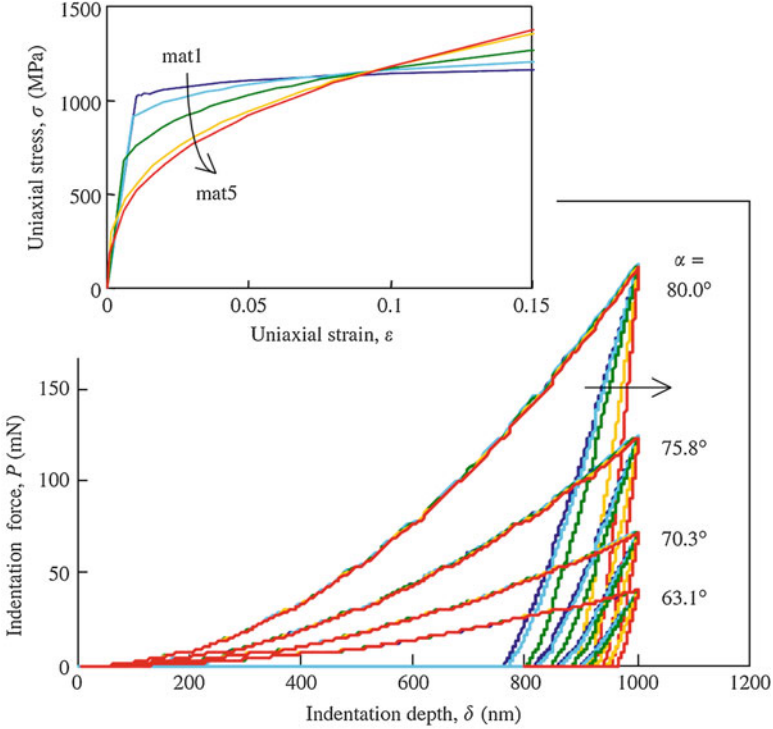


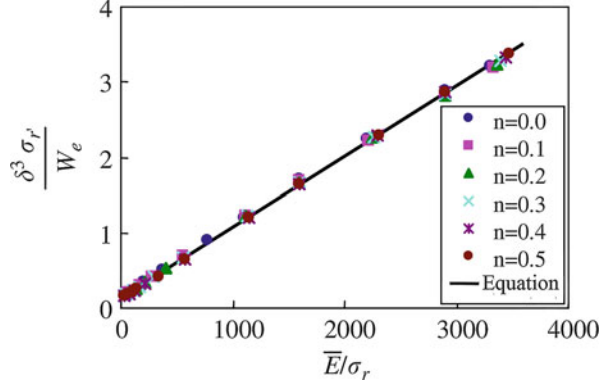
Fig. 2 A set of special materials with same loading curvatures when different indenter angles (63.14° , 70.3° , 75.79° , 80.0°) are used. The five materials have elastoplastic properties (E , σ_y , n) as mat1 (100 GPa, 1022 MPa, 0.05), mat2 (105.0GPa, 911.4 MPa, 0.10), mat3 (120.0 GPa, 678.3 MPa, 0.19), mat4 (200.0 GPa, 299.8 MPa, 0.33), and mat5 (300.0 GPa, 185.0 MPa, 0.37), respectively. The uniaxial stress-strain curves of these special materials are given in the inset on the top-left corner (Chen et al. 2007)

known, from the unloading triangle (plus the knowledge of C), the other two shape factors can be approximately derived. In order to make the best overall matching of the unloading curves, we take the normalized unloading work W_e as the governing unloading shape factor.

Since the unloading work depends on both the contact stiffness (which is related with \bar{E}) and the maximum load (which is related with C), a new representative stress for unloading, σ_r , is sought such that the normalized unloading work is related with both \bar{E} and C , but is essentially independent of n . A representative example is given in Fig. 3 for the Berkovich indenter where the unloading work is fitted by

$$\frac{\delta^3 \sigma_r}{W_e} = \lambda_1 \frac{\bar{E}}{\sigma_r} + \lambda_0 \quad (14)$$

Fig. 3 The relationship between $\delta^3 \sigma_r / W_e$ and \bar{E} / σ_r for plastic materials, with $\bar{E} / \sigma_r > 30$ or so (Chen et al. 2007)



for plastic material with $\bar{E} / \sigma_r > 30$. $\lambda_1 = 0.0009322$ and $\lambda_0 = 0.1402$ for $\alpha = 70.3^\circ$, and they take different values for other α . The representative stress σ_r is given by

$$\sigma_r = R \left(1.3 \frac{\sigma_r}{E} + 2.6 \varepsilon_R \right)^n, \tag{15}$$

which is valid for α between 60° and 80° . The simple functional forms derived in this section permit an explicit derivation of special materials with almost same unloading work, elaborated below.

Special Materials with Same Loading and Unloading Curves for a Sharp Indenter

Although Alkorta et al. (2005) have determined special materials with the same loading and unloading curves for one particular indenter, here we report an improved procedure to explicitly derive such materials, which also sets a part of the basis for finding the mystical materials. Based on Eq. 14, for a given indenter #A with half-apex angle α^A and two materials with elastoplastic properties (E_1, σ_{y1}, n_1) and (E_2, σ_{y2}, n_2) , if they are to have the same unloading work, they need to satisfy

$$\frac{\delta^3}{W_{e1}^A} = \lambda_1 \frac{\bar{E}_1}{(\sigma_{r1}^A)^2} + \lambda_0 \frac{1}{\sigma_{r1}^A} = \frac{\delta^3}{W_{e2}^A} = \lambda_1 \frac{\bar{E}_2}{(\sigma_{r2}^A)^2} + \lambda_0 \frac{1}{\sigma_{r2}^A}, \tag{16}$$

where $\varepsilon_R^A = 0.0319 \cot \alpha^A$, and $\sigma_{r1}^A = R_1 (1.3 \sigma_{r1} / E_1 + 2.6 \varepsilon_R^A)^{n_1}$ and $\sigma_{r2}^A = R_2 (1.3 \sigma_{r2} / E_2 + 2.6 \varepsilon_R^A)^{n_2}$ are the unloading representative stresses for materials #1 and #2 of indenter #A, respectively. Thus,

$$\sigma_{r2}^A = \frac{\lambda_0 + \sqrt{\lambda_0^2 + 4 \left(\lambda_1 \frac{\bar{E}_1}{(\sigma_{r1}^A)^2} + \lambda_0 \frac{1}{\sigma_{r1}^A} \right) (\lambda_1 \bar{E}_2)}}{2 \left(\lambda_1 \frac{\bar{E}_1}{(\sigma_{r1}^A)^2} + \lambda_0 \frac{1}{\sigma_{r1}^A} \right)}. \quad (17)$$

Therefore, the procedure of deriving two materials with different elastoplastic properties yet with almost the same loading and unloading curves (for indenter #A) can be concluded as

- (a) Choose any E_1 and E_2 that are different (with fixed $\nu_1 = \nu_2 = 0.3$).
- (b) Choose any value of σ_{y1} and n_1 (as long as material #1 remains sufficiently plastic), and derive $R_1 = \sigma_{y1} (E_1 / \sigma_{y1})^{n_1}$.
- (c) Calculate σ_{R1}^A from Eq. 8 and solve for σ_{R2}^A from Eq. 10.
- (d) Obtain a flow stress-total strain pair of the uniaxial stress-strain curve for material #2 as $\sigma_2^A = \sigma_{R2}^A$ and $\varepsilon_2^A = 2\varepsilon_R^A + 2\sigma_{R2}^A / E_2$.
- (e) Calculate σ_{R1}^A from Eq. 15 and solve for σ_{R2}^A from Eq. 17.
- (f) Obtain another flow stress-total strain pair for material #2 as $\sigma_2^a = \sigma_{r2}^A$ and $\varepsilon_2^a = 2.6\varepsilon_R^A + 1.3\sigma_{r2}^A / E_2$.
- (g) From the stress-strain pairs, solve $n_2 = \ln(\sigma_2^A / \sigma_2^a) / \ln(\varepsilon_2^A / \varepsilon_2^a)$, $R_2 = \sigma_2^A / (\varepsilon_2^A)^{n_2}$, and $\sigma_{y2} = \left((E_2)^{n_2 / (n_2 - 1)} / (R_2)^{1 / (n_2 - 1)} \right)$. Finally, carry out a numerical test to verify that indeed $C_1^A = C_2^A$ and $W_{e1}^A = W_{e2}^A$.

Therefore, for any material #1 with (E_1, σ_{y1}, n_1) , a special material #2 with (E_2, σ_{y2}, n_2) can be explicitly derived such that they yield indistinguishable loading and unloading curves for a given conical indenter. There are infinite sets of such special materials, and identification of these is no longer based on “trial and error.” An example of such is given in Fig. 4, where the effectiveness of the proposed approach is validated (with the difference of C and W_e less than 0.5% – such very small difference is due to the error of fitting functions and numerical solutions which is inevitable). This pair of material cannot be distinguished by only using the Berkovich indenter, yet their P - δ curves may become separable with other distinct indenter angles.

According to Eq. 16, if $\bar{E}_1 > \bar{E}_2$ it can be shown that $\sigma_{r1}^A > \sigma_{r2}^A$, whereas from Eq. 10, $\sigma_{R1}^A > \sigma_{R2}^A$. For most plastic materials, the representative strain is much larger than yield strain; thus, the two identified total strains satisfy $\varepsilon_2^a > \varepsilon_2^A$ – this implies that for a pair of special materials, if $\bar{E}_1 > \bar{E}_2$, then $n_1 > n_2$ and $\sigma_{y1} < \sigma_{y2}$; in addition, the stress-strain curves of material #1 and #2 must intersect between ε_2^a and ε_2^A . Thus, if two special materials also have the same loading and unloading curves for indenter #B, their intersection point must also be placed between ε_2^b and ε_2^B ; this is only possible when the difference between α^A and α^B is not too extreme (since both ε_R^A and ε_R^B vary as indenter angle changes, and such variation is more prominent for sharper angles). That is, the mystical materials, if they exist, should be valid within a specified range of indenter angles – more discussions are given in the next subsection.

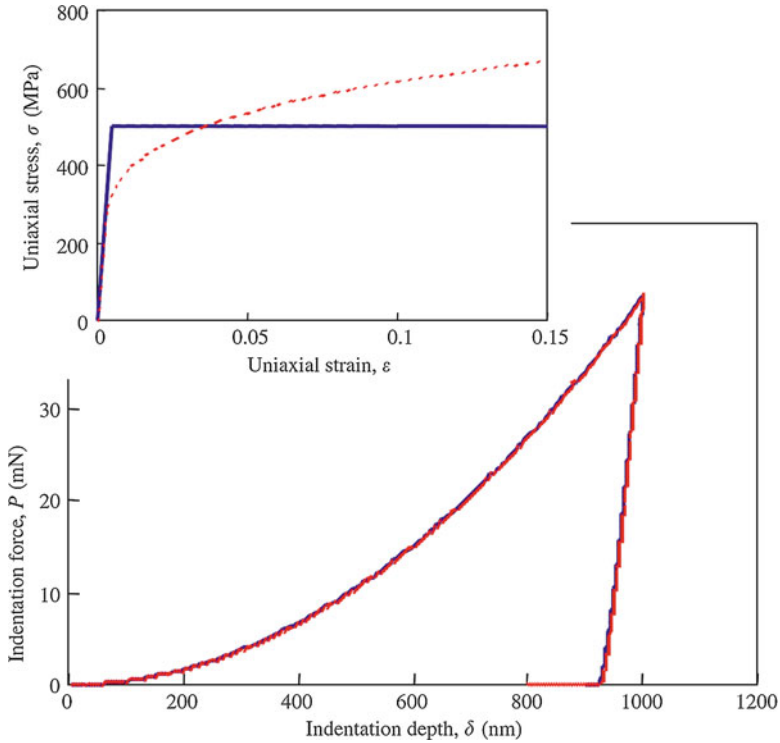


Fig. 4 A pair of special materials with indistinguishable loading and unloading curves when $\alpha=70.3$. The black solid curve represents the material with $(E, \sigma_y, n) = (100.0\text{GPa}, 500.0\text{ MPa}, 0.0)$, and the red dash curve represents the material with $(E, \sigma_y, n) = (110.0\text{GPa}, 295.0\text{ MPa}, 0.2)$. The uniaxial stress-strain curves of these special materials are given in the inset on the top-left corner (Chen et al. 2007)

Determine Mystical Materials with Same Loading and Unloading Curve for Dual Indenters

Weak-Form Mystical Materials and Their Possible Existence

In most previous dual (or plural) indentation approaches with Poisson's ratio fixed, the material elastic modulus is first obtained from the unloading curve of one particular indenter (e.g., $\alpha=70.3^\circ$). Next, the representative stress-based approach is used to determine two (or more) flow stress-total strain points on the uniaxial stress-strain curve, by utilizing the loading curvatures obtained from dual (or plural) sharp indenters. Therefore, if we could identify a pair of special materials with elastoplastic properties (E_1, σ_{y1}, n_1) and (E_2, σ_{y2}, n_2) (their Poisson's ratio $\nu_1 = \nu_2 = 0.3$), such that they yield indistinguishable loading and unloading P - δ curves for indenter #A (e.g., $\alpha=70.3^\circ$), and also almost identical loading curvature

for another indenter #B, then the conventional dual (or plural) indentation analysis would fail since it cannot promise unique solution. This pair of special materials, which does not require their unloading works to match for indenter #B, may be termed as the weak-form mystical materials.

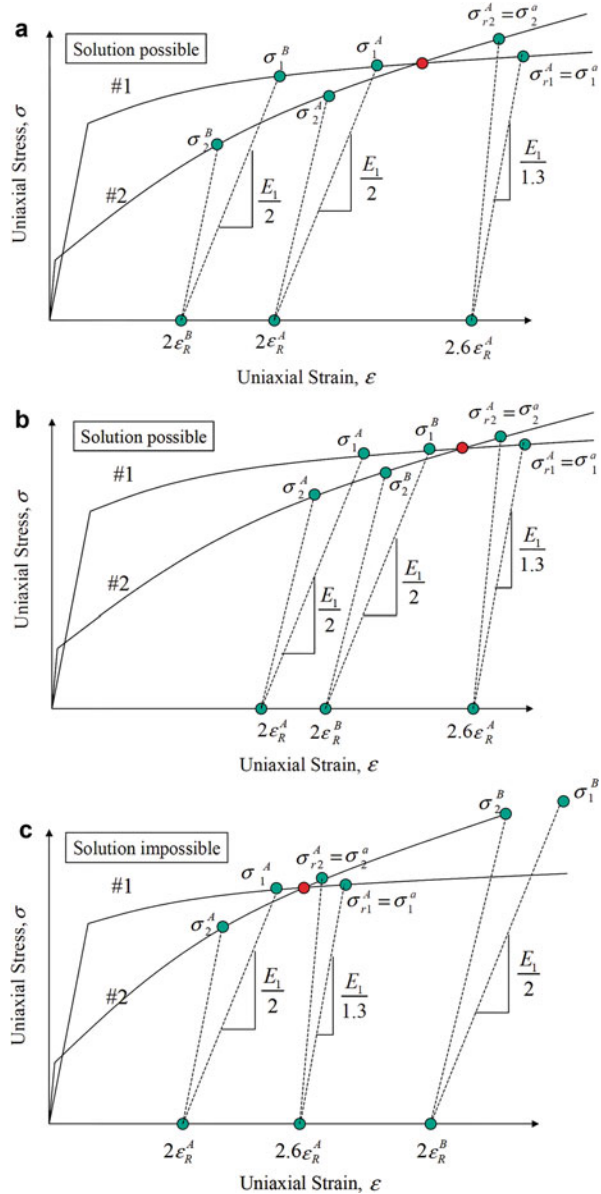
Meanwhile note that in a displacement-controlled experiment where the maximum penetration is fixed, if $C_1^A = C_2^A$ and $C_1^B = C_2^B$, then the maximum indentation load for these two materials are also the same. Thus, to make their unloading works to match for indenter #A ($W_{e1}^A = W_{e2}^A$), their contact stiffness must be fairly close ($S_1^A \approx S_2^A$), the two materials must have close Young's moduli ($\bar{E}_1 \approx \bar{E}_2$), and this also implies their unloading works for indenter #B must also be very close ($W_{e1}^B \approx W_{e2}^B$). Thus, the weak-form mystical materials are very close to the mystical materials we are looking for.

What general properties must the weak-form mystical materials satisfy (assuming they are sufficiently plastic such that Eq. 14 applies)? Since these materials must be a subset of the special materials derived from the above procedures, therefore, with ($\alpha^B > \alpha^A$), if ($\bar{E}_1 < \bar{E}_2$), the two mystical materials must satisfy $n_1 < n_2$ and $\sigma_{y1} > \sigma_{y2}$; moreover, the uniaxial stress-strain curves of these two materials must intersect outside $2\varepsilon_R^A$ but inside $2.6\varepsilon_R^A$. A schematic showing of the relative status of the two mystical materials is given in Fig. 5a. Since these two candidates need to intersect within a relatively small region, this implies that the mystical materials would only exist for a specified range of indenter angle (with $\alpha^B > \alpha^A$) and material properties. For the current case, the difference between α^B and α^A cannot be too extreme, and the materials need to be sufficiently plastic (with large E_1/σ_{y1} and E_2/σ_{y2}) so as to leave enough possible space for materials #1 and #2 to intersect within the desired region.

Next, we consider the case with $\alpha^B < \alpha^A$, but the difference is not so much such that $2\varepsilon_R^A < 2\varepsilon_R^B < 2.6\varepsilon_R^A$. If $\bar{E}_1 < \bar{E}_2$, then from the above procedures, $\sigma_{R1}^A = \sigma_1^A > \sigma_{R2}^A = \sigma_2^A$ and $\sigma_{R1}^B = \sigma_1^B > \sigma_{R2}^B = \sigma_2^B$; since $m_p^A/m_e^A > m_p^B/m_e^B$, so the difference between σ_1^A and σ_2^A is larger than that of σ_1^B and σ_2^B ; moreover, the unloading representative stresses satisfy $\sigma_{r1}^A = \sigma_1^a > \sigma_{r2}^A = \sigma_2^a$. All these features lead to $n_1 < n_2$ and $\sigma_{y1} > \sigma_{y2}$, and a possible solution for the mystical pair #1 and #2 is sketched in Fig. 5b. In this case, the uniaxial stress-strain curves of these two materials must intersect between $2\varepsilon_R^B$ and $2.6\varepsilon_R^A$, and such range is even narrower than that in Fig. 5a. That is, the mystical materials may exist in a small space of material properties and indenter angles; nevertheless, such solution is possible.

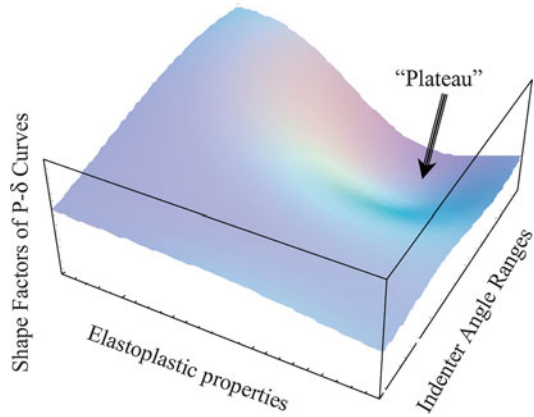
When α^B is much smaller than α^A such that $2\varepsilon_R^B > 2.6\varepsilon_R^A$, all the relative magnitudes of the representative stresses discussed in the last paragraph still hold, except that now those related with $2\varepsilon_R^B$ are moved to the right side of those related with $2.6\varepsilon_R^A$ – according to the new schematic in Fig. 5c, it is impossible to find a solution for the mystical material. This implies that if the indenter #A is the Berkovich tip, the indenter #B must be larger than about 64.88° such that the mystical materials may exist according to Fig. 5a, b. Similarly, if #B is taken to be the Berkovich tip (since the weak-form mystical material is very close to the desired mystical material), then #A must be smaller than 74.50° . Therefore, rigorously

Fig. 5 Schematics of possible solutions of a pair of weak-form plastic mystical materials; assuming $\bar{E}_1 < \bar{E}_2$, then $n_1 < n_2$ and $\sigma_{y1} > \sigma_{y2}$. The indenter angle $A(\alpha^A)$ is fixed (which equals to 70.3° in many existing dual indenter techniques). **(a)** When $\alpha^B > \alpha^A$ solution is possible. **(b)** When $\alpha^B < \alpha^A$ but $2\varepsilon_R^B < 2.6\varepsilon_R^A$, solution is possible. **(c)** When $\alpha^B < \alpha^A$ and the difference between these two angles is large such that $2\varepsilon_R^B > 2.6\varepsilon_R^A$, rigorous solution is not possible (Chen et al. 2007)



speaking, under the premise that Poisson’s ratio is always fixed at 0.3, if a Berkovich tip is used in the dual indenter method, the mystical materials would only be possible if the other indenter angle is between 64.88° and 74.50° . Moreover, such mystical materials need to be relatively plastic.

Fig. 6 The schematic of the possible existence of the mystical materials and relevant indenter angle ranges – they correspond to the plateaus on the multidimensional surface of the shape factors of the P - δ curves (Chen et al. 2007)



In a short summary, there is no pair of mystical materials that would be applicable to arbitrary large range of indenter angles, and only for limited material space-indenter angle combinations can the mystical siblings be found – this is part of the reason the mystical materials were not discovered in the past. The current problem of finding the mystical materials is analogous to a multivariable problem (Fig. 6), where the indentation shape factors (the general z-axis) are related with the material elastoplastic properties (the general x-axis) and indenter angle ranges (the general y-axis) through a multidimensional surface. On such surface, the slope may not be large and distinct everywhere, and small regions that are relatively flat may exist, as sketched in Fig. 6. The mystical materials and the relevant indenter angle ranges would correspond to such plateaus (with zero or almost zero slope) of the multidimensional surface – although these plateaus are small compared with the entire parameter space, they may still exist and thus have important theoretical value for probing the uniqueness of indentation analysis. The search of such plateaus is elaborated next.

Search for Mystical Materials with Fixed Poisson's Ratio

In most literature of indentation analysis, Poisson's ratio was fixed by neglecting its influence. In order to challenge the basis of the uniqueness solution and be consistent with the literature, a fixed $\nu = 0.3$ is used in most part of this chapter.

For a given pair of indenters #A and #B, and a given material #1 which is sufficiently plastic, in order to identify the weak-form mystical siblings, Eqs. 10, 11, and 17 need to be satisfied rigorously, from which σ_{R2}^A , σ_{R2}^B , and σ_{r2}^A can be derived (with any specified \bar{E}_2), respectively. Unfortunately, the only solution that rigorously satisfies all these equations is exactly material #1. On the other hand, note that there are no perfect finite element solution and numerical fitting, and all numerical results are subject to small error. It is then acceptable if small errors are added to fitting functions Eqs. 10, 11, and 17, such that the resulting

loading/unloading P - δ curves of material #2 would be indistinguishable to that of material #1, with the difference between their shape factors below several percent – such small perturbation is also inevitable in the data measured from any real experiment. Another advantage is that, when small perturbation to the shape function is allowed, the existence range of the mystical material is significantly enlarged, in terms of both the material property and indenter angle space. In essence, although it is difficult to search for perfect plateaus (with exactly zero slope) on the multidimensional surface describing the indentation shape factors, it is possible to search for plateaus with slopes that are very close to zero through a numerical algorithm – the results are still the indistinguishable mystical materials given the inevitable small perturbations in numerical and experimental indentation tests. The search process is also relatively straightforward since the simple and explicit formulations of the primary shape factors of loading/unloading P - δ curves are established earlier. Of course, with Poisson's ratio fixed at 0.3, such identification procedure is no longer explicit.

The numerical search process is the following (with fixed $\nu_1 = \nu_2 = 0.3$):

- (a) Choose any E_1 and α^A (e.g., the Berkovich tip).
- (b) Choose any initial value of α^B and then iterate such that its difference with respect to α^A is increased.
- (c) Choose initial values of n_1 and σ_{y1} and then iterate, preferably in the plastic material range.
- (d) Choose an initial value of E_2 which is at least 5% different than E_1 and then iterate, such that the difference could become bigger. This ensures that the initial guesses of material #1 and #2 are sufficiently different.
- (e) From Eqs. 10, 11, and 17, solve for σ_{R2}^A , σ_{R2}^B , and σ_{r2}^A .
- (f) Give small errors (few percent) to Eqs. 10, 11, and 17. For example, if we wish to increase σ_{R2}^A by ω times (e.g., $\omega=1.01$ for a 1% error), then from Eq. 6 the new representative stress of material #2 that is related with loading for indenter #A becomes

$$\sigma_{R2-error}^A = \left(\frac{1}{\omega \sigma_{R2}^A} + \frac{m_p^A}{m_e^A} \frac{1}{E_2} \left(\frac{1}{\omega} - 1 \right) \right)^{-1} \quad (18)$$

Similarly, when error is permitted, the two other representative stresses are

$$\sigma_{R2-error}^B = \left(\frac{1}{\omega \sigma_{R2}^B} + \frac{m_p^B}{m_e^B} \frac{1}{E_2} \left(\frac{1}{\omega} - 1 \right) \right)^{-1} \quad (19)$$

and

$$\sigma_{R2-error}^A = \left(\lambda_0 + \sqrt{\lambda_0^2 + 4 \left(\frac{\lambda_1}{\omega} \frac{\bar{E}_2}{(\sigma_{r2}^A)^2} + \frac{\lambda_0}{\omega} \frac{1}{\sigma_{r2}^A} \right) (\lambda_1 \bar{E}_2)} \right) / \left(2 \frac{\lambda_1}{\omega} \frac{\bar{E}_2}{(\sigma_{r2}^A)^2} + 2 \frac{\lambda_0}{\omega} \frac{1}{\sigma_{r2}^A} \right), \quad (20)$$

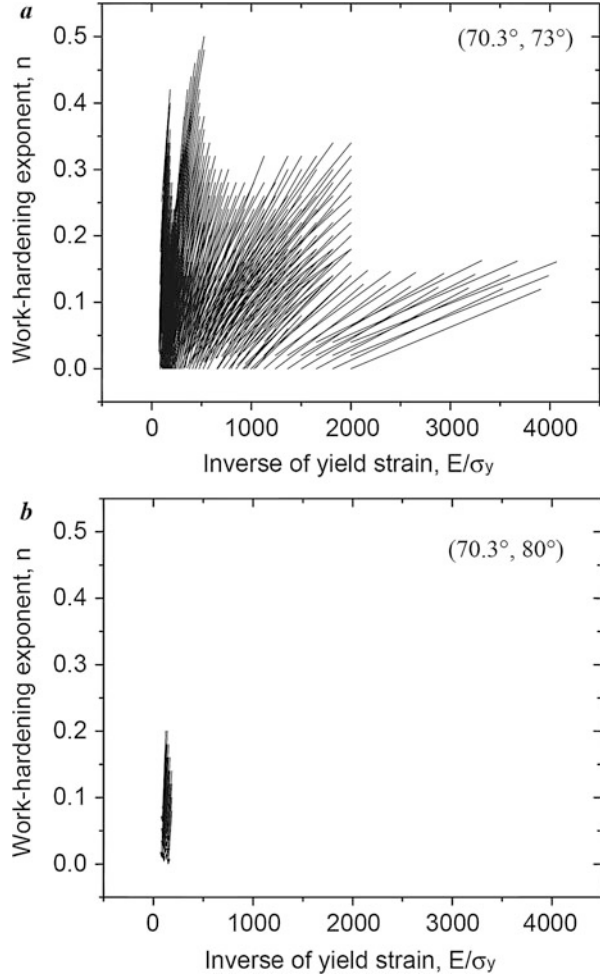
respectively. Therefore, by plus or minus several percent of error, the upper and lower bounds of the error bars of σ_{R2}^A , σ_{R2}^B , and σ_{r2}^A can be derived – when the three error bars are combined, an error band can be formed.

- (g) For all possible combinations of σ_{R2}^A , σ_{R2}^B , and σ_{r2}^A within the error band, the admissible solutions of material #2 are sought; since σ_{R2}^A , σ_{R2}^B , and σ_{r2}^A and E_2 must satisfy certain compatibility, only a small portion of the combinations within the error band could become candidate materials. For any admissible solution, their loading curvatures for indenters #A and #B can be estimated from Eq. 6, and their unloading work for indenter #A is obtained from Eq. 14; the results are then compared with that of material #1, and the more promising pairs with smaller errors are recorded along with the current indenter angle range. A candidate pair of mystical materials is found if the computed error is smaller than 2% for all shape factors.
- (h) Iterate E_2 .
- (i) Iterate n_1 and σ_{y1} .
- (j) Iterate α^B .
- (k) Lastly, numerical indentation analyses are performed on the most promising candidate mystical material pairs using finite element simulations, to confirm that their loading and unloading curves are visually indistinguishable (i.e., the shape factors of their indentation curves are sufficiently close).

The existence range of the mystical materials is first explored through a series of maps in terms of materials space ($E/\sigma_y, n$). In Fig. 7a, b, the pairs of identified mystical materials are shown in line segments – both ends of each segment represent two mystical materials with different elastoplastic properties, yet they yield almost indistinguishable loading/unloading P - δ curves. For a given line segment, any sets of materials along the length of the segment are also mystical materials, and the most distinct mystical materials can be found at the ends of the longest segment. The area where the density of the segments is large indicates a possible gold mine of mystical materials.

For any given indenter angle range, the mystical materials can only exist within a certain region. As the difference between the dual indenters becomes larger, the existence range of the mystical material becomes smaller; moreover, the segments are shorter which also indicates the differences between their elastoplastic properties are smaller. Therefore, if a pair of extreme indenter angles (very sharp and very blunt) is used, the mystical materials do not exist (however, in experiments, the

Fig. 7 The existence maps of mystical materials for given indenter angle ranges, $(\alpha^A, \alpha^B) =$ (a) $(70.3^\circ, 73^\circ)$, (b) $(70.3^\circ, 80^\circ)$. Each segment links a pair of mystical materials within the material space $(E/\sigma_y, n)$. A gold mine of mystical materials is discovered when E/σ_y is about 100 and when n is small (Chen et al. 2007)



use of a pair of extreme sharp indenters is often impractical). In the examples illustrated in Fig. 7, $\alpha^A = 70.3^\circ$ is always used because many previous studies rely on measuring the elastic modulus from a Berkovich indentation.

When the indenter angle range is relatively small, e.g., $(\alpha^A, \alpha^B) = (70.3^\circ, 73^\circ)$ (Fig. 7a), the mystical materials can be found in a large range, but notably for materials with smaller n . Quite a few of the more plastic mystical materials can exist with large differences of their elastoplastic properties. No mystical materials are available with both large n and large E/σ_y .

By contrary, when the indenter angle difference becomes larger, e.g., $(\alpha^A, \alpha^B) = (70.3^\circ, 80^\circ)$ (Fig. 7b), then the survival range of mystical materials are confined to the lower-left corner of the materials space map. Mystical materials become possible only for materials with $n = 0-0.2$ and $E/\sigma_y \approx 100$. Note that

if a pair of mystical materials is identified for a larger angle range, they are still mystical siblings for any subrange of indenter angles. Moreover, once a map for $(\alpha^A, \alpha^B) = (70.3^\circ, 60^\circ)$ is made and combined with Fig. 7b, their common elements are the extreme mystical materials that are effective when $(\alpha^A, \alpha^B) = (60^\circ, 80^\circ)$. Therefore, the materials with small n and with E/σ_y around 100 represent the gold mine around which many mystical materials can be identified. In fact, there are quite a few important engineering metals and alloys near this area, for example, Ti alloys, Ni alloys, Mg alloys, and high strength steel, in addition to a few ceramics and polymers (Ashby 1999) – extra care is needed for the measurement of their elastoplastic properties.

For the mystical materials identified in this section, with $\nu = 0.3$ but without knowing other information in advance, within the specified dual indenter angle range, their elastoplastic properties cannot be uniquely determined from the indentation analysis since their loading and unloading curves are indistinguishable.

Numerical indentation tests are carried out on these materials, and their P - δ curves are compared, which also reveal the characteristics of the mystical materials.

In Fig. 8a, for the indenter angle range $(\alpha^A, \alpha^B) = (70.3^\circ, 74^\circ)$, a pair of plastic mystical materials is chosen, with $E_1 = 100$ GPa, $\sigma_{y1} = 50$ MPa, $n_1 = 0.06$, and $E_2 = 110$ GPa, $\sigma_{y2} = 29.336$ MPa, and $n_2 = 0.17277$, respectively. Their uniaxial stress-strain curves are shown in the inset, and their corresponding indentation load-displacement curves are given. It is apparent that their indentation behaviors are almost identical (for both loading and unloading curves); specifically, the difference between C_1^A and C_2^A is about 2%, the difference between C_1^B and C_2^B is about 1%, and the difference between W_{e1}^A and W_{e2}^A is about 2%. The difference between W_{e1}^B and W_{e2}^B is about 4%; nevertheless, during the search of weak-form mystical materials, the matching criterion is not applied to the unloading curves with indenter #B. Note that when a larger indenter angle range is used, such as 63.14° and 75.79° , their P - δ curves become quite separable, and therefore it is still possible to use the established dual indenter method to measure the elastoplastic properties of these two materials with the wider indenter angle ranges.

In fact, a wider indenter angle separation also means that the two identified total strains are further apart, which also gives better numerical accuracy – a rule of thumb is that the two identified total strain should be separated by at least 30%, which is qualified for most plastic materials with indenter angles 63.14° and 75.79° (Ogasawara et al. 2007b), including the example in Fig. 8a. On the other hand, for the more elastic materials, the separation between the identified total strain points becomes smaller for a given indenter angle range. To ensure accuracy, a large indenter angle range $\Delta\alpha = |\alpha^A - \alpha^B|$ is recommended for more elastic materials with $E/\sigma_y < 100$ or so.

Figure 8b gives an intriguing example, which corresponds to $E_1 = 100$ GPa, $\sigma_{y1} = 872.47$ MPa, $n_1 = 0.0$, and $E_2 = 103.75$ GPa, $\sigma_{y2} = 715.61$ MPa, and $n_2 = 0.10663$. This pair of extreme mystical material leads to almost the same indentation loading/unloading behaviors when the indenter angle changes from 60° to 80° , which has reached the limit of the sharp indenter angles used in this study. Apparently, many existing dual (or plural) indentation methods in the literature

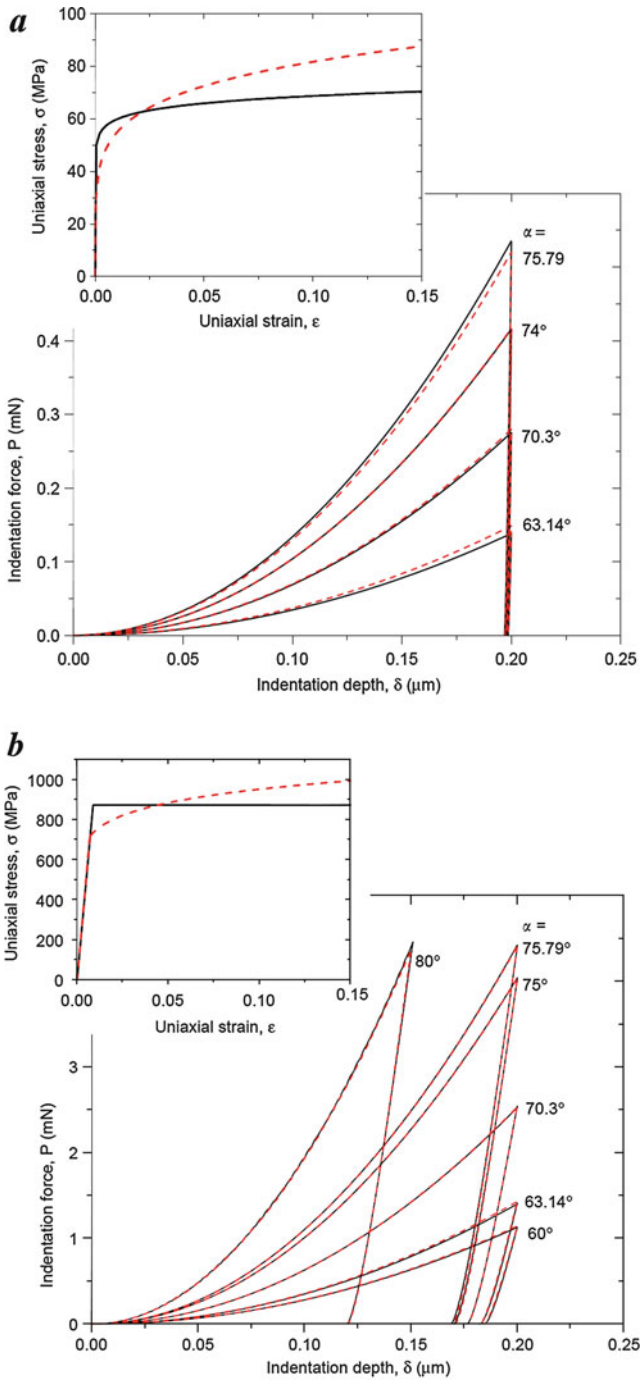


Fig. 8 (continued)

would fail to distinguish them, and more extreme indenters are needed. However, this is often not practical because without advanced knowledge, one could not predetermine what kind of $\Delta\alpha$ needs to be used in an experiment. Moreover, when extreme indenter angles are used, the measured hardness would differ by orders of magnitude, and new problems such as those associated with the resolution of the instrument, size effect, indenter tip alignment, and indentation cracking will emerge. In the next section, we will introduce how to use alternative methods to distinguish such extreme mystical pair.

For a pair of mystical materials near or inside the gold mine, their stress-strain curves tend to intersect around a total strain of 0.05. Moreover, if one material has a larger plane-strain modulus, then it always has larger work-hardening exponent and smaller yield stress.

Alternative Methods to Distinguish Mystical Materials

Improved Spherical Indentation

Spherical indenter has the unique advantage that with one penetration, the loading curvature is reduced as if the indenter angle becomes smaller. Since the mystical materials may be eventually distinguishable by a large $\Delta\alpha$, we only need to control the δ_{\max} during spherical indentation. For that matter, the penetration depth has to be sufficiently deep and a few previous studies (e.g., Cao and Lu 2004) where $\delta_{\max}/r = 0.1$ do not qualify; indeed, if δ_{\max}/r is small (such as 0.1), the spherical indenter method still cannot distinguish the mystical materials because the effective $\Delta\alpha$ is small. Alternatively, in one of our works (Zhao et al. 2006b), an improved spherical indentation technique was proposed, which seems promising since $\delta_{\max}/r = 0.3$, which mimics a very sharp indenter angle.

Figure 9a shows the spherical indentation result on the extreme mystical material pair derived from Fig. 8b. Initially, when the penetration is shallow, which is analogous to the blunter indenter angles, the two P - δ curves cannot be separated. However, when δ/r is larger than about 0.15, these two materials become distinguishable. Finally, at the maximum penetration, there is about 8% difference between their C . Although these two materials have very close contact stiffness and also C measured at $\delta/r = 0.13$, the relatively large difference at

Fig. 8 Representative case studies of mystical materials: the P - δ curves of mystical materials and the uniaxial σ - ε curves of mystical materials are given in inset on the top-left corner. **(a)** A pair of plastic mystical materials for $(\alpha^A, \alpha^B) = (70.3^\circ, 74^\circ)$, with $E_1 = 100$ GPa, $\sigma_{y1} = 50$ MPa, $n_1 = 0.06$ (solid curve), and $E_2 = 110$ GPa, $\sigma_{y2} = 29.336$ MPa, $n_2 = 0.17277$ (dash curve). They can be distinguished when more different indenter angles are used. **(b)** A pair of extreme mystical materials with $E_1 = 100$ GPa, $\sigma_{y1} = 872.47$ MPa, $n_1 = 0.0$ (solid curve), and $E_2 = 103.75$ GPa, $\sigma_{y2} = 715.61$ MPa, $n_2 = 0.10663$ (dash curve). They cannot be effectively distinguished by indenter angles from 60° and 80° (Chen et al. 2007)

$\delta_{\max}/r = 0.3$ is sufficient to make the spherical indentation technique work well. By following the reverse analysis procedure in Zhao et al. (2006b), the determined values are $E_1 = 102.5$ GPa, $\sigma_{y1} = 828.85$ MPa, $n_1 = 0.024$, and $E_2 = 106.34$ GPa, $\sigma_{y2} = 701.30$ MPa, and $n_2 = 0.108$, respectively. All errors (except that for n_1 cannot be counted since its true value is 0.0) are smaller than about 2% (except σ_{y1} which is about 5% and still within reasonable range). Finally, the identified uniaxial stress-strain curve from the reverse analysis of the improved spherical indentation method (Zhao et al. 2006b) is given in the inset, and its excellent capability of distinguishing the mystical material is justified. The improved spherical indentation technique proposed by Zhao et al. (2006b) only requires one simple indentation test, and thus it is convenient and reliable. This technique is therefore recommended for materials inside or near the gold mine of mystical materials.

Film Indentation

When an elastoplastic film with finite thickness is bonded to a rigid substrate, the increased conical penetration dramatically increases the loading curvature as if the sharp indenter angle is increased. Thus, the substrate effect provides an alternative way of obtaining extreme indenter angles to distinguish the mystical materials. We have proposed a theory, where the loading curvatures at the penetration of 1/3 and 2/3 of film thickness (h), along with the unloading work, are used to obtain the material elastoplastic properties from one film indentation test (Zhao et al. 2006a). When this technique is applied to the extreme mystical materials (found in Fig. 8b, in Fig. 9b), it can be readily seen that their loading curvatures become quite different at $\delta_{\max}/h = 2/3$, which is again sufficient to distinguish the extreme mystical pair although their loading curvatures at $\delta/r = 1/3$ are close. In addition, their unloading works are also different.

By following the reverse analysis described in Zhao et al. (2006a), finally the determined properties of the extreme pair of mystical materials are $E_1 = 97.5$ GPa, $\sigma_{y1} = 872.47$ MPa, $n_1 = 0.0$ and $E_2 = 103.75$ GPa, $\sigma_{y2} = 719.19$ MPa, $n_2 = 0.105$, respectively. All errors (except that for n_1 cannot be counted since its true value is 0.0) are smaller than 1.5% except E_1 which is only -2.5% . The identified uniaxial stress-strain curve from the film indentation technique (Zhao et al. 2006a) is given in the inset of Fig. 10, and it is demonstrated to be able to distinguish the mystical materials with high accuracy. Note that the advantage of the film indentation technique is that it may be applied to specimens with finite thickness; however, it also requires the testing platform (i.e., the substrate) to be sufficiently stiff and hard, which may not be practical in some cases (see Zhao et al. (2006a) for discussions). Nevertheless, it could be used as an alternative method to distinguish the mystical materials and is proven to work well.

In fact, from the error analysis of both the improved spherical indentation and film indentation methods (Zhao et al. 2006a, b), both techniques have the best accuracy for materials inside or near the gold mine of mystical materials, which make them complementary to the dual (or plural) sharp indentation analysis.

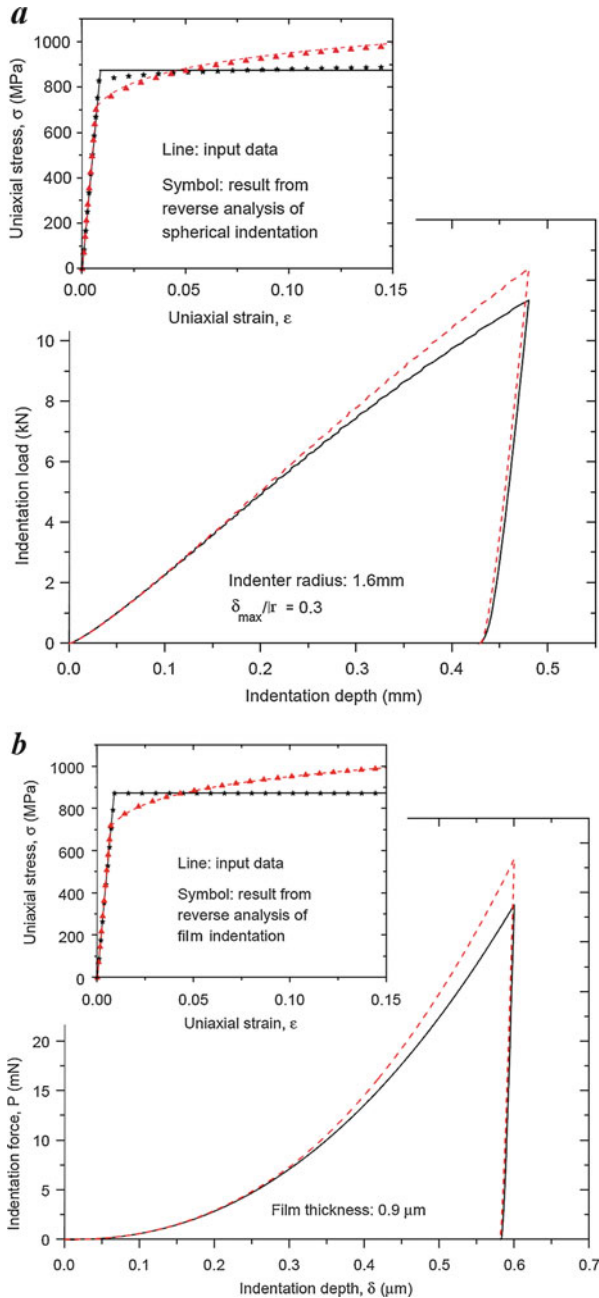


Fig. 9 The P - δ curves of two alternative indentation on the extreme mystical materials found in Fig. 8b. (a) Improved spherical indentation method (Zhao et al. 2006b). (b) Film indentation method (Zhao et al. 2006a). These methods can distinguish the extreme mystical materials – in the inset on top-left corner, the uniaxial stress-strain curves obtained from reverse analysis (symbols) are compared with the input (true) data and show excellent agreement (Chen et al. 2007)

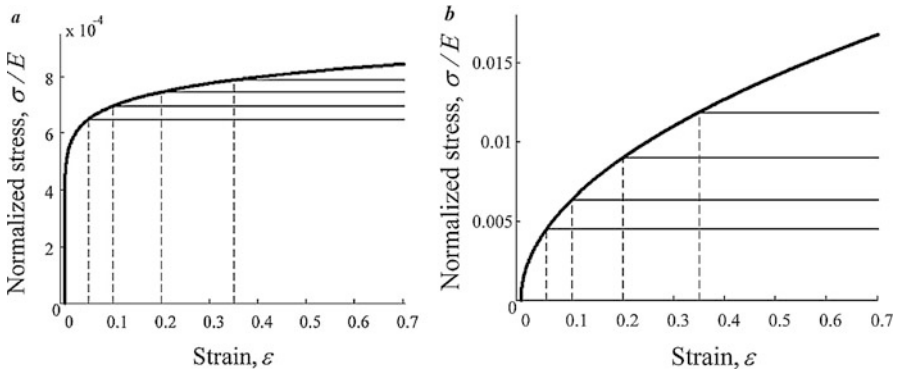


Fig. 10 Materials tailoring of two power-law materials (a) $E/\sigma_y = 2500$ and $n = 0.1$ and (b) $E/\sigma_y = 2500$ and $n = 0.5$. Thick lines are the uniaxial true stress-strain curves of untailed materials, and each thin line presents a modified material. Every modified material is defined as initially having the same constitutive curve with the original power-law material but becomes perfectly plastic beyond a bifurcation strain, ε_b . In the examples in both (a) and (b), ε_b equals 0.05, 0.1, 0.2, and 0.35 for the four tailored materials (Liu et al. 2009)

Detectable Strain Range of Indentation Test

The Critical Strain

When a sharp indenter penetrates a bulk specimen, the loading curvature C is only a function of the material elastoplastic; theoretically, any variation of the material constitutive relationship (the stress-strain curve) can, to a certain degree, cause C to deviate from its original value. However, there exists a critical strain beyond which tailoring material properties will no longer induce prominent variation to the indentation response (e.g., less than 1% deviation in the measured C). [The 1% threshold is set because it is a typical order-of-magnitude intrinsic error associated with numerical indentation analyses: below this critical level, one can hardly tell whether the difference is caused by material tailoring or by numerical error. Of course it could set a different threshold (e.g., 0.5%), but the critical strains derived from the new threshold are not going to be much different from the ones identified in this study, and the relevant conclusions still hold.] Therefore, indentation is limited to probing material elastoplastic properties within a particular strain range below critical, and the tailoring of the stress-strain curve beyond this range cannot be detected by indentation reverse analysis, leading to a non-unique solution.

To verify the existence of the critical strain and identify its dependence on material properties and indenter angles, without losing generality, we consider six representative power-law materials with $E/\sigma_y = 2500, 1000,$ and 100 and $n = 0.1$ and 0.5 . Two example stress-strain curves are given as thick lines in Fig. 10a, b. To tailor the constitutive relationship, one could choose any bifurcation strain $\varepsilon_b > \varepsilon_y$

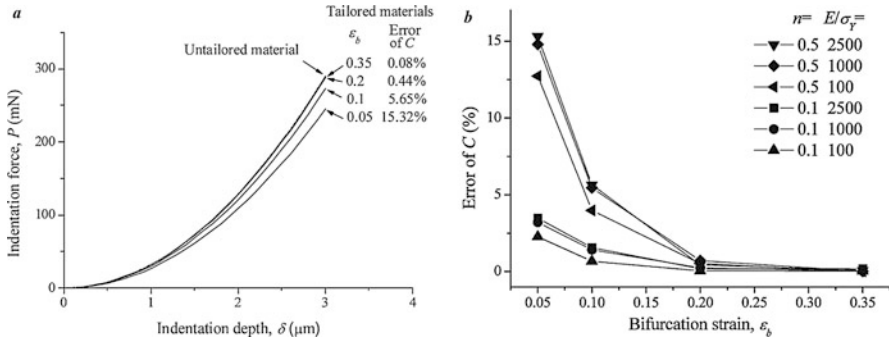


Fig. 11 (a) Indentation loading curves (with $\alpha=70.3^\circ$) for a representative power-law material and four artificial materials tailored at different ϵ_b . The error of C is calculated as the difference of loading curvatures between each tailored materials and the untailored material. (b) The evolution of the error of C (with $\alpha=70.3^\circ$) as a function of ϵ_b induced by material tailoring. A critical strain ϵ_c can be defined when the error of C falls below 1% (Liu et al. 2009)

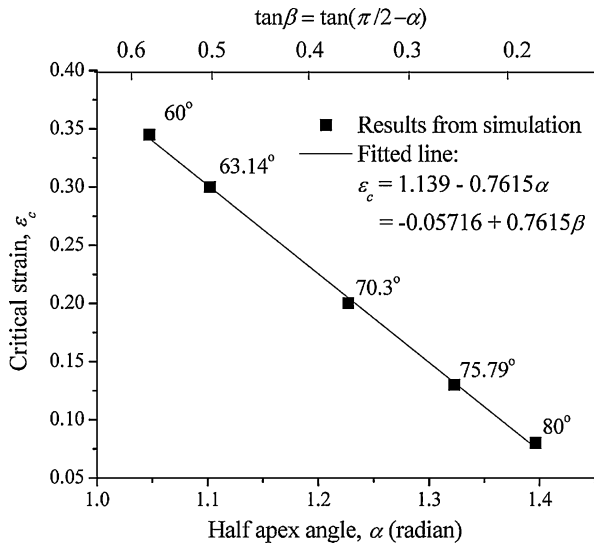
and assign perfect plastic behavior after this point (the flat thin line in Fig. 10). In other words, the tailored stress-strain relationship becomes

$$\begin{aligned}
 \sigma &= E\varepsilon, \text{ for } \varepsilon \leq \varepsilon_y \\
 \sigma &= R\varepsilon^n, \text{ for } \varepsilon_y \leq \varepsilon \leq \varepsilon_b \\
 \sigma &= R\varepsilon_b^n = \text{constant}, \text{ for } \varepsilon \geq \varepsilon_b.
 \end{aligned}
 \tag{21}$$

Such a tailoring strategy provides a rough upper bound of the moderate modifications of the hardening function. For the selected cases shown in Fig. 10, ϵ_b equals 0.05, 0.1, 0.2, and 0.35 for the four modified materials. Upon indentation, the modified material will in principle show a different C from that of the original (unmodified) material, but S remains essentially unchanged since the elastic modulus is unaffected; thus, we only focus on the perturbation of C during loading.

Numerical indentation tests are carried out on the original power-law materials as well as their modified counterparts (with ϵ_b varying over a large range and example load-displacement curves given in Fig. 11a). The indenter angle α is also varied. The percentage error of C (between the original and modified materials) can be plotted as a function of ϵ_b , given in Fig. 11b for a representative indenter angle $\alpha = 70.3^\circ$. Each line in this Figure represents modifications of one of the six original materials, and each symbol in that line denotes the percentage error of C for a tailored material characterized by a particular ϵ_b . The error decreases quickly with ϵ_b , since a higher ϵ_b means that the constitutive relationship of the modified material is closer to the original one (Fig. 10). In addition, the error of C is smaller when n is smaller or when E/σ_y is smaller (and the effect of n is more prominent), and this is also related to the fact that the difference between the stress-strain curves of the original and tailored materials are smaller when n and/or E/σ_y is smaller (Fig. 10 and Eq. 21).

Fig. 12 The dependency of the critical strain ϵ_c on the half-apex angle, α , of sharp indenters. Also shown is the linear fitting Eq. 22 (Liu et al. 2009)



The most important finding is, regardless of the original material, the percentage error of C tends to be lower than the 1% threshold when ϵ_b exceeds a critical value, ϵ_c . For $\alpha = 70.3^\circ$, the critical strain ϵ_c is identified as 0.20 – any modification of the plastic behavior beyond this point cannot be effectively reflected on the P - δ curve, and the modified material would exhibit an indistinguishable indentation response with respect to the original material as long as $\epsilon_b > \epsilon_c$; under this circumstance, both the original and modified materials are possible solutions of the reverse analysis of the same P - δ curve. Note that the difference between the constitutive relationships of the original material and modified material can be substantial especially when n is large and/or when the critical strain is small.

Using the sharp indentation technique, the stress-strain curve may only be probed when the strain is between 0 and ϵ_c (the detectable strain range or sometimes the detectable range for short). When α is fixed, ϵ_c is essentially material-independent, and this is verified by numerical analyses using a wide range of materials with diverse elastoplastic properties (as long as ϵ_y is not too large, which is satisfied for most metals and alloys). (Although we derived the critical strain by modifying the power-law material model in this section, the approach can be extended to other material models, and we have verified that the value of critical strain is not sensitive to the material model used.) In Fig. 12, the critical strain ϵ_c is presented as a function of the indenter angle, where the relationship is almost linear and can be fitted as

$$\epsilon_c = 1.139 - 0.7615\alpha \tag{22}$$

within the indenter angle range in this chapter (where α is in radians in Eq. 22). It is interesting that the detectable range of sharp indenters can be small for blunt indenters, and one of the popular indenters with $\alpha = 80^\circ$ could only prove strain

up to about 8% and thus is sensitive to the non-uniqueness issue; the use of sharper indenters could partially reduce this concern. However, sharp indenters may cause cracking, and the results may be sensitive to friction in practice. In addition, even the sharpest indenter used in this study ($\alpha = 60^\circ$) could only detect up to 35% of strain, and this performance falls well below common expectations. We therefore conclude that it is impossible to measure the entire stress-strain curve uniquely via a sharp indentation test. (Although the plural indenter technique, especially those with sharper indenters (e.g., using $\alpha = 60^\circ$ and 70.3°), could alleviate the non-uniqueness problem than those with blunter indenters (e.g., using $\alpha = 80^\circ$ and $\alpha = 70.3^\circ$), none of them would work well outside the critical strain range.)

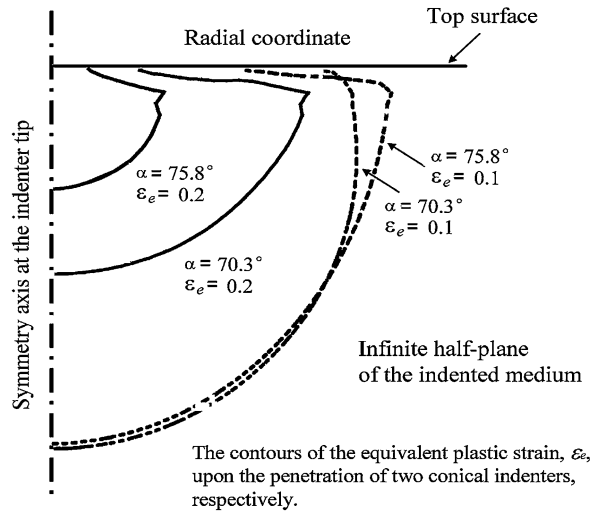
Besides challenging the uniqueness of indentation test, the discovery of the critical strain has several other impacts. First, during the verification of an established indentation method, often the stress-strain curve of a real engineering material needs to be fitted into power-law form and serve as a benchmark for examination (Guelorget et al. 2007); however, a different fitting range of the stress-strain curve would lead to different fitted results of the plastic parameters, and the fitting range should be consistent with the detectable range of indentation test, otherwise a large bias would occur (Ogasawara et al. 2008). Second, the critical strain could also guide numerical indentation analyses. From either the stress-strain curve measured in a lab experiment or with respect to a specific material model, a data set of the uniaxial stress and strain is needed as the input for material properties to be used in an FEM program/simulation. Sometimes there is a concern as to how many data points are needed and how refined they need to be. Here we show that regardless of the details, only the input stress-strain data within the detectable strain range is relevant, and outside this range, the data is essentially unimportant (in terms of the resulting P - δ curves).

Variation of Critical Strain: A Qualitative Explanation

A qualitative explanation of the critical strain, along with its dependency on the indenter geometry, lies in the nonuniform plastic deformation below the indenter. Since the work done by the indenter during the penetration of a sharp indenter is $W_t = C\delta^3/3$, the perturbation of C may be understood from that of W_t , which equals the total deformation energy. The material deformation energy includes two parts: the recoverable strain energy and the plastic dissipation; the later part equals $(W_t - W_e)$.

The influence of tailoring the material constitutive behavior is nonuniform in the indented solid. Within the field of equivalent plastic strain (ε_e) produced by indentation, only the regions with $\varepsilon_e > \varepsilon_b$ are more sensitive to the modification of plastic properties beyond ε_b . Figure 13 shows the contour plots of ε_e in the deformed unmodified power-law solid (with $E/\sigma_y = 2500$, and $n = 0.5$), and the contour lines of $\varepsilon_e = 0.1$ and 0.2 are given when $\alpha = 70.3^\circ$ and 75.79° , respectively. For fair comparison, the contours shown in Fig. 13 are taken at the instants when the indenters have done the same amount of work. The material enclosed roughly by

Fig. 13 The contour plot of the equivalent plastic strain, ε_e , in a semi-infinite solid indented by two sharp indenters. The superimposed contours are taken from independent tests at the instants when the indenters have done the same amount of work (Liu et al. 2009)



the solid contour lines (e.g., $\varepsilon_e = 0.2$) may be perturbed if the stress-strain curve is modified beyond the corresponding value, for example, $\varepsilon_b = 0.2$. Within the area where $\varepsilon_e < \varepsilon_b$, if the deformation energy is relatively small, material tailoring would not yield prominent variation of the overall indentation response.

For $\alpha = 70.3^\circ$ for example, the deformation energy in the region enclosed by the contour of $\varepsilon_e = 0.2$ is only about 22% of the energy enclosed by the contour of $\varepsilon_e = 0.1$. Therefore, tailoring the material beyond $\varepsilon_b = 0.1$ will include a larger variation to the overall P - δ curve than tailoring above $\varepsilon_b = 0.2$. This is qualitatively consistent with the descending trend of the error of C due to tailoring Fig. 11b. When the indenter angle is varied, for $\alpha = 75.79^\circ$, the fraction of deformation energy of the region enclosed by the contour of $\varepsilon_e = 0.2$ is only 3% of its counterpart for $\alpha = 70.3^\circ$, and the fraction of deformation energy of the region surrounded by the contour of $\varepsilon_e = 0.1$ is about half that for $\alpha = 70.3^\circ$. This suggests that material tailoring beyond the same strain ($\varepsilon_b = 0.1$ or 0.2) could influence C more significantly for the sharper indenter, and thus qualitatively speaking, ε_c is larger for a sharper indenter and smaller for a blunter indenter.

Conclusion

Although indentation tests have been long used to measure the elastoplastic properties of engineering materials, a systematic study on the uniqueness of indentation analysis, i.e., on the possible existence of the one-to-one correspondence between the indentation load-displacement curves, material parameters, and indenter geometries, is still lacking. Among the available indentation techniques, the dual (or plural) sharp indenter method is often considered as well established, and it is also the foundation of many other similar indentation analyses and has been widely

used in practice. In this chapter, through a comprehensive numerical study, the primary shape factors of the indentation load-displacement curves are related with the material properties and indenter angles through simple functional forms. Both explicit and numerical procedures are established to search for mystical materials with distinct elastoplastic properties yet yield indistinguishable load-displacement curves, even when the indenter angles are varied. Consequently, these mystical materials cannot be distinguished by many of the existing dual (or plural) indenter methods or spherical indenter methods (if the indentation depth is shallow). The properties and the existence of such mystical materials are discussed.

In addition, a critical strain is identified as the upper bound of the detectable range of indentation, and moderate tailoring of the constitutive behavior beyond this range cannot be effectively detected by the reverse analysis of the load-displacement curve. That is, for a given indenter geometry, beyond the critical strain, there is no unique solution of the material plastic behavior from the reverse analysis of the load-displacement curve. The critical strain is identified as a function of the sharp indenter angle, through which the analysis of sharp indentation may be qualitatively correlated to some extent – this link also enriches the indentation theory and applications.

The topics in this chapter address the important question of the uniqueness of indentation test, as well as providing useful guidelines to properly use the indentation technique to measure material elastoplastic properties.

Acknowledgments The work is supported in part by National Science Foundation CMS-0407743 and CMMI-CAREER-0643726 and in part by the Department of Civil Engineering and Engineering Mechanics, Columbia University.

References

- J. Alkorta, J.M. Martinez-Esnaola, J.G. Sevillano, Absence of one-to-one correspondence between elastoplastic properties and sharp-indentation load–penetration data. *J. Mater. Res.* **20**, 432–437 (2005)
- M.F. Ashby, *Materials Selection in Mechanical Design*, 2nd edn. (Elsevier, Amsterdam, 1999)
- F.P. Bowden, D. Tabor, *The Friction and Lubrications of Solids* (Oxford University Press, Oxford, 1950)
- Y.P. Cao, J. Lu, A new method to extract the plastic properties of metal materials from an instrumented spherical indentation loading curve. *Acta Mater.* **52**, 4023–4032 (2004)
- T.W. Capehart, Y.T. Cheng, Determining constitutive models from conical indentation: sensitivity analysis. *J. Mater. Res.* **18**, 827–832 (2003)
- X. Chen, N. Ogasawara, M. Zhao, N. Chiba, On the uniqueness of measuring elastoplastic properties from indentation: the indistinguishable mystical materials. *J. Mech. Phys. Solids* **55**, 1618–1660 (2007)
- Y.T. Cheng, C.M. Cheng, Can stress–strain relationships be obtained from indentation curves using conical and pyramidal indenters? *J. Mater. Res.* **14**, 3493–3496 (1999)
- Y.T. Cheng, C.M. Cheng, Scaling, dimensional analysis, and indentation measurements. *Mater. Sci. Eng.* **R44**, 91–149 (2004)
- B. Guelorget, M. Francois, C. Liu, J. Lu, Extracting the plastic properties of metal materials from microindentation tests: experimental comparison of recently published methods. *J. Mater. Res.* **22**, 1512–1519 (2007)

- L. Liu, N. Ogasawara, N. Chiba, X. Chen, Can indentation technique measure unique elastoplastic properties? *J. Mater. Res.* **24**, 784–800 (2009)
- S.D. Mesarovic, N.A. Fleck, Spherical indentation of elastic-plastic solids. *Proc. R. Soc. Lond. A* **455**, 2707–2728 (1999)
- N. Ogasawara, N. Chiba, X. Chen, Representative strain of indentation analysis. *J. Mater. Res.* **20**, 2225–2234 (2005)
- N. Ogasawara, N. Chiba, X. Chen, Limit analysis-based approach to determine the material plastic properties with conical indentation. *J. Mater. Res.* **21**, 947–958 (2006)
- N. Ogasawara, N. Chiba, M. Zhao, X. Chen, Measuring material plastic properties with optimized representative strain-based indentation technique. *J. Solid Mech. Mater. Eng.* **1**, 895–906 (2007a)
- N. Ogasawara, N. Chiba, M. Zhao, X. Chen, Comments on “Further investigation on the definition of the representative strain in conical indentation” by Y. Cao and N. Huber [*J. Mater. Res.* **21**, 1810 (2006)]: A systematic study on applying the representative strains to extract plastic properties through one conical indentation test. *J. Mater. Res.* **22**, 858–868 (2007b)
- N. Ogasawara, M. Zhao, N. Chiba, X. Chen, Comments on “Extracting the plastic properties of metal materials from microindentation tests: experimental comparison of recently published methods” by B. Guelorget, et al. [*J. Mater. Soc.* **22**, 1512 (2007)]: The correct methods of analyzing experimental data and reverse analysis of indentation tests. *J. Mater. Res.* **23**, 598–608 (2008)
- I.N. Sneddon, The relationship between load and penetration in the axisymmetric Boussinesq problem for a punch of arbitrary profile. *Int. J. Eng. Sci.* **3**, 47–57 (1965)
- K.K. Tho, S. Swaddiwudhipong, Z.S. Liu, K. Zeng, J. Hua, Uniqueness of reverse analysis from conical indentation tests. *J. Mater. Res.* **19**, 2498–2502 (2004)
- M. Zhao, X. Chen, N. Ogasawara, A.C. Razvan, N. Chiba, D. Lee, Y.X. Gan, A new sharp indentation method of measuring the elastic-plastic properties of soft and compliant materials by using the substrate effect. *J. Mater. Res.* **21**, 3134–3151 (2006a)
- M. Zhao, N. Ogasawara, N. Chiba, X. Chen, A new approach of measuring the elastic-plastic properties of bulk materials with spherical indentation. *Acta Mater.* **54**, 23–32 (2006b)



Helical Buckling Behaviors of the Nanowire/Substrate System

7

Youlong Chen, Yilun Liu, and Xi Chen

Contents

Introduction	242
Helical Buckling Mechanism for a Stiff Nanowire on the Surface of an Elastomeric Substrate	244
Experimental	244
Theoretical Analysis	245
FEM Simulations	256
Mechanism of the Transition from In-plane Buckling to Helical Buckling for a Stiff Nanowire on the Surface of an Elastomeric Substrate	266
Model and Method	266
Results and Discussion	268
Helical Buckling of a Nanowire Embedded in a Soft Matrix Under Axial Compression	275
Theoretical Analysis	276
Results and Discussion	279
Conclusion	284
References	285

Y. Chen (✉)

International Center for Applied Mechanics, State Key Laboratory for Strength and Vibration of Mechanical Structures, School of Aerospace, Xi'an Jiaotong University, Xi'an, China
e-mail: cy1900125@126.com

Y. Liu

International Center for Applied Mechanics, State Key Laboratory for Strength and Vibration of Mechanical Structures, School of Aerospace, Xi'an Jiaotong University, Xi'an, China
e-mail: yilunliu@mail.xjtu.edu.cn

X. Chen

Department of Earth and Environmental Engineering, Columbia Nanomechanics Research Center, Columbia University, New York, NY, USA
e-mail: xichen@columbia.edu

Abstract

When a nanowire is deposited on a compliant soft substrate or embedded in matrix, it may buckle into a helical coil form when the system is compressed. Using theoretical and finite element method (FEM) analyses, the detailed three-dimensional coil buckling mechanism for a silicon nanowire (SiNW) on a poly-dimethylsiloxane (PDMS) substrate is discussed. A continuum mechanics approach based on the minimization of the strain energy in the SiNW and elastomeric substrate is developed, and the helical buckling spacing and amplitude are deduced, taking into account the influences of the elastic properties and dimensions of SiNWs. These features are verified by systematic FEM simulations and parallel experiments. When the debonding of SiNW from the surface of the substrate is considered, the buckling profile of the nanowire can be divided into three regimes, i.e., the in-plane buckling, the disordered buckling in the out-of-plane direction, and the helical buckling, depending on the debonding density. For a nanowire embedded in matrix, the buckled profile is almost perfectly circular in the axial direction; with increasing compression, the buckling spacing decreases almost linearly, while the amplitude scales with the $1/2$ power of the compressive strain; the transition strain from 2D mode to 3D helical mode decreases with the Young's modulus of the wire and approaches to 1.25% when the modulus is high enough, which is much smaller than nanowires on the surface of substrates. The study may shed useful insights on the design and optimization of high-performance stretchable electronics and 3D complex nanostructures.

Keywords

Buckling mode · Nanowire · Soft substrate · Helical mode · In-plane mode · Transition · Embedded wire · Continuum mechanics · FEM · Post-buckling behaviors · Buckling wavelength · Buckling amplitude

Introduction

Recently, stretchable electronics has attracted wide research interests and holds great potential applications, such as precision metrology (Stafford et al. 2004; Wilder et al. 2006), electronic eye cameras (Ko et al. 2008; Rogers et al. 2010), flexible displays (Chen et al. 2002; Crawford 2005), stretchable electronic circuits (Kim et al. 2008; Song et al. 2009), and conformable skin sensors (Lacour et al. 2005; Someya et al. 2004), to name a few. In stretchable electronics, the fragile and stiff elements (e.g., silicon, metal films, or wires) are usually placed on the elastomeric substrates and precompressed to some fundamental buckling modes (Audoly and Boudaoud 2008; Charlot et al. 2008; Chen and Yin 2010). The buckling conformation of the brittle and stiff elements can provide large deformability and tolerance for stretching, compression, bending, twisting, and even combined loading modes of the stretchable electronics (Rogers et al. 2010).

In essence, through buckling the compressive strain is released and replaced by relative small bending strain of the slender compressed structure. In this way, the brittle component of electronics can be fabricated into complex buckled form by precisely adjusting the geometrical structures of the stiff components and the elastomeric substrates, the mechanical properties of every constitutive component, and the adhesion between the stiff components and the substrates (Ko et al. 2008; Xu et al. 2015). Further reduction of the thickness or radius of the buckling members, although beneficial for accommodating more compressibility (Song et al. 2009; Xiao et al. 2010), is unfortunately limited by the fabrication processes and functionality of the flexible electronics. Besides, the bending strain is localized at the crest and valley of the buckling configurations for sinusoidal buckling such as out-of-plane buckling of nanowires and films. Hence, the failure of the brittle components usually initiates at the stress concentration region. Furthermore, the electronic properties of silicon components are closely dependent on the strain applied in the components, and the localized strain may cause nonuniform electronic properties in the components (Peng et al. 2009; Sajjad and Alam 2009). An alternative solution to further enhancing the deformability of the buckling mode is to make the buckles go three-dimensional, such as the helical buckling model of nanowire that has been successfully applied to silicon nanowire (SiNW) which can sustain very large stretchability up to the failure strain of poly-dimethylsiloxane (PDMS) (Xu et al. 2011). This helical buckling mode can easily handle diverse loading models, including multi-axial stretching, compression, bending, and twisting, and extend the usefulness of stretchable electronic elements. Furthermore, systems at micro- or even nanoscales were fabricated (Chen and Yin 2010), and coaxial electrospinning with the helical configuration and other spring-like structures were obtained based on buckling of nanofibers and films on curved substrates (Chen and Yin 2010; Yin and Chen 2010).

The buckling of nanowires on elastomeric substrate is a very common phenomenon and has important applications in stretchable electronics (Durham and Zhu 2013; Kim et al. 2008; Ryu et al. 2009; Wang et al. 2013), and different buckling structures have been reported. For example, the buckled nanowire could lie within the plane of the substrate (in-plane buckling), or perpendicular to the substrate (out-of-plane buckling), or of special interest here is the helical configuration (combination of the in-plane and out-of-plane modes) (Xiao et al. 2008, 2010; Xu et al. 2011). The adhesion strength between SiNW and elastomeric substrate may have an important role in regulating the buckling mode of SiNW on PDMS substrate. It has been reported that the ultraviolet/ozone (UVO) treatment may strengthen the interaction between SiNW and PDMS by forming strong covalent bonds, and without the treatment, there only exist much weaker Van der Waals forces (Efimenko et al. 2002; Qin and Zhu 2011). Consequently, no debonding between SiNW and PDMS is observed after proper UVO treatment and SiNW buckles in the helical mode. While sliding at the interface is distinctly detected without UVO (with low adhesion strength) and hence SiNW shows the in-plane buckling mode with lower strain energy to the out-of-plane buckling.

Buckling behaviors of slender wires under compression have been studied for centuries, and buckling modes are strongly affected by the boundary conditions of the wire. For a beam without continuous support, such as a cantilever or a simply supported beam, the Euler buckling theory is fairly effective to describe the buckling configuration and critical loads (Goriely et al. 2008; Oldfather et al. 1933; Timoshenko and Gere 2009; Zeeman 1976). Although the initial curved nanoribbons with selecting adhesion to the elastomeric substrate can also generate complex three-dimensional buckle modes, the strain distribution along the nanowires are also not uniform and the loose contact between the nanoribbons and the substrate may initiate failure (Xu et al. 2015), and the selecting adhesion is more challenging than perfect (uniform) adhesion.

Meanwhile, when embedded in a soft matrix with low Young's modulus, both 2D sinusoidal (planar) and 3D helical (nonplanar) buckling modes of compressed wires were observed, while a stiffer matrix would drive wires to buckle only in the 3D mode (Su et al. 2014). Based on the Winkler foundation model (Timoshenko and Gere 2009), they developed the 2D buckling theory of a wire embedded in a soft matrix and concentrated on the critical strain and initial buckling wavelength (or wavevector) of sinusoidal buckled wires (Slesarenko and Rudykh 2016; Su et al. 2014), but the 3D helical buckling mechanism and post-buckling behaviors of embedded wires have not been explicitly explained.

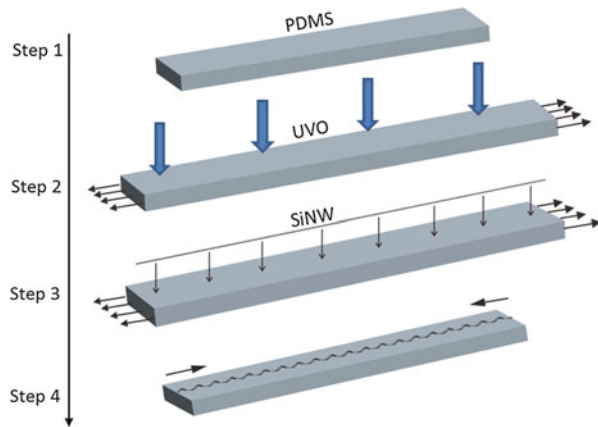
In this chapter, we first focus on studying the intrinsic helical buckling mechanism of a straight SiNW on the surface of PDMS substrate with perfect interface adhesion via theoretical analysis and comprehensive finite element method (FEM) simulations, and further verifying by parallel experiments. A continuum mechanics approach is established which is extendable to three-dimensional helical coil buckling on elastomeric substrates. Then, the compressive buckling of a nanowire partially bonded to an elastomeric substrate is studied via finite element method (FEM) simulations and experiments, to explore the mechanism of the transition from in-plane buckling to helical buckling. Thereafter, the helical buckling mechanics is extended to buckling behaviors of embedded wires in an infinite matrix, and the post-buckling of post-buckling evolution of embedded wires is studied.

Helical Buckling Mechanism for a Stiff Nanowire on the Surface of an Elastomeric Substrate

Experimental

The experimental method previously developed (Xu et al. 2011) was employed in this work. The experimental results are shown in section “[Helical Buckling of a Nanowire Embedded in a Soft Matrix Under Axial Compression](#)” for comparison with analytical and FEM simulation results. Below, a brief summary of the experimental method is provided. SiNWs were synthesized on the silicon wafer by chemical vapor deposition using gold nanoclusters as catalysts and silane (SiH_4) as vapor-phase reactant. A poly(dimethylsiloxane) (PDMS) substrate with a thickness

Fig. 1 Schematic diagram of the helical buckling of SiNW on PDMS substrate (Xu et al. 2011). First, the PDMS is prestretched and radiated by ultraviolet/ozone (Step 2), which improves the adhesion strength between PDMS and SiNW. Then, the SiNW is transferred to the surface of PDMS (Step 3) using contact printing. After releasing the prestrain in PDMS, helical buckling occurs in the SiNW (Step 4)



of 2 mm was prepared using Sylgard 184 (Dow Corning) by mixing the “base” and the “curing agent” with a ratio of 10:1. The mixture was first placed in a vacuum oven to remove air bubbles and then thermally cured at 65 °C for 12 h. Rectangular slabs of suitable sizes were cut from the cured piece.

Figure 1 schematically shows the process for fabricating buckled SiNWs. A miniaturized tensile testing stage (Ernest F. Fullam) was used to mechanically stretch the PDMS slab to the desired levels of prestrain, with both ends of the slab clamped. The prestrained substrate was radiated under a UV lamp (low-pressure mercury lamp, BHK) (Fig. 1, Step 2). A contact printing method was used to transfer the SiNWs on the silicon wafer to the PDMS substrate. The silicon wafer was slid along the prestrained direction to align the SiNWs (Fig. 1, Step 3). Releasing the prestrain in PDMS resulted in buckling of the SiNWs (Fig. 1, Step 4). The releasing step was carried out in-situ under an atomic force microscope (AFM). At a number of intermittent strain levels, the releasing was paused and AFM images were taken, based on which the buckling spacing of the SiNWs is deduced. Between such intermittent strain levels, the releasing (unloading) strain rate was between $10^{-4}/s$ and $10^{-3}/s$ to eliminate the effect of loading rate on the buckling behaviors. The entire experiment was strain controlled.

Theoretical Analysis

Continuum Model for Helical Buckling Mode

Subjected to an effective compressive strain ε_{com} imposed by contraction of the PDMS substrate (from Steps 3 to 4 in Fig. 1), the SiNW buckles into a helical buckling configuration, which is schematically shown in Fig. 2. Here, the x axis is along the axial direction of the SiNW, y axis is perpendicular to the surface of PDMS, and z axis is along lateral direction in PDMS surface. Based on experimental observation (Xu et al. 2011), the helical mode of the SiNW is uniform in the middle.

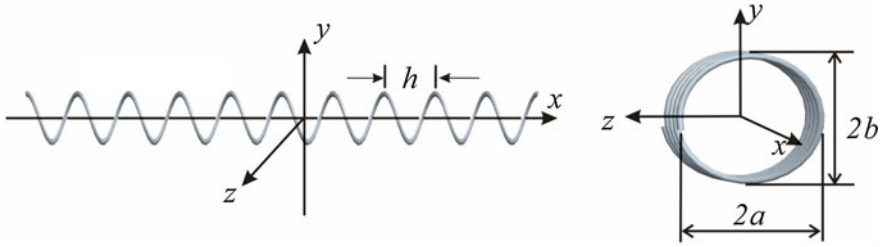


Fig. 2 Schematic diagram of the three-dimensional helical buckling of the SiNW with y direction perpendicular to the surface of PDMS substrate and the profile of the buckled SiNW

Therefore, in theoretical model the deflection of the SiNW is approximated by a helical curve as $w = a\cos(kx)$ in y direction and $v = b\sin(kx)$ in z direction, where a and b denote the deflection amplitudes in y and z directions, respectively, and k is the buckling wavevector, a variable related to the screw pitch (or buckling spacing) h by $k = 2\pi/h$. For the simplification of the theoretical analysis, the constitutive relations of the SiNW and PDMS substrate are assumed as linear elasticity, which is also consistent with the previous literatures (Mei et al. 2011; Song et al. 2009; Su et al. 2014).

Let u denote the axial displacement of the SiNW, and the membrane strain ε_m in the SiNW (caused by axial tensile or compressive force in the SiNW) is expressed as

$$\varepsilon_m = \frac{du}{dx} + \frac{1}{2} \left[\left(\frac{dv}{dx} \right)^2 + \left(\frac{dw}{dx} \right)^2 \right]. \quad (1)$$

As the Young's modulus of the SiNW (e.g., 200 GPa) is several orders of magnitude larger than the elastic modulus of the PDMS substrate (e.g., 6 MPa), the membrane stress in SiNW is assumed as a constant and the shear stress along the axial direction of SiNW is ignored in the theoretical model. Indeed, based on the FEM simulation results (see Fig. 16) the membrane strain in SiNW is almost constant and is much smaller than the bending strain. The shear stress along the axial direction of SiNW is about one order of magnitude smaller than the traction stresses perpendicular to the SiNW. However, if the Young's modulus of the wire is comparable to that of the substrate, the shear stress may play an important role in the buckling behaviors of a wire on an elastomeric substrate. The shear stress effect will be systematically explored in our future works. By assuming constant membrane strain ε_m in SiNW, Eq. 1 can be rewritten as

$$\frac{d\varepsilon_m}{dx} = \frac{d^2u}{dx^2} + \frac{dv}{dx} \frac{d^2v}{dx^2} + \frac{dw}{dx} \frac{d^2w}{dx^2} = 0. \quad (2)$$

Substituting the helix deflection of the SiNW into Eq. 2, the governing equation of the axial displacement u is

$$\frac{d^2u}{dx^2} + (a^2 - b^2)k^3 \sin(kx) \cos(kx) = 0. \quad (3)$$

Solving Eq. 3 to obtain the axial displacement

$$u = \frac{(a^2 - b^2)k}{8} \sin(2kx) + C_1x + C_2, \quad (4)$$

where the two parameters C_1 and C_2 can be determined by the boundary conditions. By ignoring the rigid body displacement of the SiNW, we obtain $C_2 = 0$. As the SiNW is perfectly bonded to the PDMS substrate, by considering the displacement compatibility between the SiNW and PDMS substrate we obtain $C_2 = -\varepsilon_{\text{com}}$, where ε_{com} is the effective compressive strain applied to the PDMS substrate. Then, the membrane strain ε_m can be obtained by substituting Eq. 4 into Eq. 1 as

$$\varepsilon_m = \frac{k^2(a^2 + b^2)}{4} - \varepsilon_{\text{com}}, \quad (5)$$

and the membrane energy (per unit length) in the SiNW is

$$U_m = \frac{k}{2\pi} \int_0^{2\pi/k} \frac{1}{2} EA \varepsilon_m^2 dx = \frac{EA}{2} \left[\frac{k^2(a^2 + b^2)}{4} - \varepsilon_{\text{com}} \right]^2, \quad (6)$$

where E and A denote the Young's modulus and cross-sectional area of the SiNW, respectively.

The curvature of the SiNW under helical mode is

$$\kappa = \frac{[a^2k^4 \cos^2(kx) + b^2k^4 \sin^2(kx) + a^2b^2k^6]^{\frac{1}{2}}}{[a^2k^2 \sin^2(kx) + b^2k^2 \cos^2(kx) + 1]^{\frac{3}{2}}}. \quad (7)$$

Then, the bending energy (per unit length) in the SiNW is

$$U_b = \frac{k}{2\pi} \int_0^{2\pi/k} \frac{EI}{2} \kappa^2 dx, \quad (8)$$

where I is the moment of inertia of the SiNW in the corresponding bending direction of the helical curve. In this work, the cross section of SiNW is circular, so that the moment of inertia $I = \pi R^4/4$ is equal in all directions, where R is the radius of SiNW. In fact, the theoretical analysis presented herein is not restricted to the

circular cross section of the SiNW. Nevertheless, in practice helical buckling is most likely to appear in nanowires with equal moment of inertia in all directions, e.g., circular cross section. Otherwise, the buckling would occur with respect to the direction with the smallest moment of inertia. By substituting Eq. 7 in to Eq. 8 and integrating, the bending energy per length is

$$U_b = \frac{EI k^4 (3a^4 k^2 + 2a^2 b^2 k^2 + 3b^4 k^2 + 4a^2 + 4b^2)}{16 (a^2 k^2 + a^2 k^4 b^2 + b^2 k^2 + 1)^{3/2}}. \quad (9)$$

For the helical buckled SiNW, additional torsion energy must be taken into account. As an approximation, the SiNW are regarded as a spring loaded by an axial force $F = EA\varepsilon_m$

$$F = EA \left[\frac{k^2 (a^2 + b^2)}{4} - \varepsilon_{com} \right]. \quad (10)$$

Then, the shear force F_S and torsional moment T acting on the cross section of the SiNW are

$$F_S = F \sqrt{\frac{(dv/dx)^2 + (dw/dx)^2}{1 + (dv/dx)^2 + (dw/dx)^2}} \quad (11)$$

and

$$T = \frac{F (w dv/dx - v dw/dx)}{\sqrt{1 + (dv/dx)^2 + (dw/dx)^2}}. \quad (12)$$

The torsion and shear energy (per unit length) in the SiNW can be given as

$$U_t = \frac{k}{2\pi} \int_0^{2\pi/k} \left(\frac{T^2}{2GI_p} + \frac{F_S^2}{2GA} \right) dx, \quad (13)$$

employing the helix deflection of the SiNW and substituting Eqs. 10–12 into Eq. 13, the torsional energy per unit length is

$$U_t = \frac{AE^2 (a^2 k^2 + b^2 k^2 - 4\varepsilon_{com})^2 \left(Aa^2 b^2 k^2 + I_p \left(\sqrt{(1+b^2 k^2)(1+a^2 k^2)} - 1 \right) \right)}{32GI_p \sqrt{(1+b^2 k^2)(1+a^2 k^2)}}, \quad (14)$$

where G is the shear modulus and $I_p = \pi R^4/2$ the polar moment of inertia.

In order to derive the deformation of the PDMS substrate caused by the helical buckling of SiNW, the helical buckling is decomposed into two buckling modes,

i.e., the in-plane buckling mode and out-of-plane buckling mode. Then, the deformation of the PDMS substrate is the superposition of displacement fields caused by the two buckling modes, respectively. As the thickness of PDMS substrate is much larger than the deflection of SiNW, the substrate is regarded as a semi-infinite solid. Based on the beam theory, the lateral distributed load (per unit length) of a beam can be derived by the deflection v and axial force $F = EA\varepsilon_m$ of the beam as $P = EId^4v/dx^4 - EA\varepsilon_md^2v/dx^2$. For the helical buckling mode, the deflection can be decomposed into out-of-plane deflection $w = a\cos(kx)$ and in-plane deflections $v = b\sin(kx)$, respectively. Then, the distributed load applied on the SiNW in y direction (out-of-plane direction) is $T_y = EId^4w/dx^4 - EA\varepsilon_md^2w/dx^2 = -P_y\cos(kx)$ where $P_y = -EAAk^2(a^2k^2/4 - \varepsilon_{com}) - Elak^4$, while the distributed load in z direction is $T_z = EId^4v/dx^4 - EA\varepsilon_md^2v/dx^2 = -P_z\sin(kx)$ where $P_z = -EAbk^2(b^2k^2/4 - \varepsilon_{com}) - Elbk^4$.

Using the Green's function method, for unit normal force acting at point $(x_1, 0, z_1)$ on the surface of an incompressible semi-infinite solid, the normal displacement at point $(x, 0, z)$ can be given as $[(x - x_1)^2 + (z - z_1)^2]^{-1/2}/(\pi\bar{E}_S)$, where $\bar{E}_S = E_S/(1 - \nu_S^2)$ is the plane-strain modulus of the elastomeric substrate with ν_S the Poisson's ratio of the substrate. For the average normal force $P_y\cos(kx)/(2R)$ over the width $2R$, the normal displacement on the surface of the PDMS substrate can be integrated as $w_{sub} = \int_{-R}^R \frac{P_y\cos(kx)}{\pi\bar{E}_SR} K_0(k|z - z_1|) dz_1$, where $K_0(k|z - z_1|)$ is the modified Bessel function of the second kind (Abramowitz and Stegun 1972; Timoshenko et al. 1970). As the buckling spacing of SiNW is much larger than the radius of SiNW, i.e., $kR \ll 1$, the dominant term in the Taylor series expansion of the normal displacement is

$$w_{sub} = \frac{P_y\cos(kx)}{\pi\bar{E}_SR} [2R(1-\gamma + \ln 2) - (R+z)\ln(k|R+z|) - (R-z)\ln(k|R-z|)], \quad (15)$$

where $\gamma = 0.577$ is the Euler's constant. Following the same procedure, for unit lateral force (z direction) acting at the point $(x_1, 0, z_1)$ on the surface of an incompressible semi-infinite solid, the lateral displacement at point $(x, 0, z)$ is $\frac{(1-\nu_S)(x-x_1)^2+(z-z_1)^2}{\pi(1-\nu_S)\bar{E}_S[(x-x_1)^2+(z-z_1)^2]^{3/2}}$. Then, for the average lateral force $P_z\sin(kx)/(2R)$ over the width $2R$, the lateral displacement (z direction) on the surface of the PDMS substrate can be integrated as $v_{sub} = \int_{-R}^R \frac{P_z\sin(kx)}{\pi\bar{E}_SR(1-\nu_S)} [(1-\nu_S)K_0(k|z - z_1|) + \nu_S k|z - z_1|K_1(k|z - z_1|)] dz_1$, where $K_1(k|z - z_1|)$ is the modified Bessel function of the second kind. For $kR \ll 1$, the dominant term in the Taylor series expansion of the lateral displacement is

$$v_{sub} = \frac{P_z\sin(kx)}{\pi R\bar{E}_S} \left[2R \left(\frac{1}{1-\nu_S} - \gamma + \ln 2 \right) - (R+z)\ln(k|R+z|) - (R-z)\ln(k|R-z|) \right] \quad (16)$$

Therefore, the strain energy of the PDMS substrates (unit length) caused by the out-of-plane and in-plane buckling of the SiNW is

$$\begin{aligned}
 U_S &= U_{\text{out-of-plane}} + U_{\text{in-plane}} = \frac{k}{2\pi} \int_0^{2\pi/k} \int_{-R}^R \frac{1}{2} \left[\frac{P_y \cos(kx)}{2R} w_{\text{sub}} + \frac{P_z \sin(kx)}{2R} v_{\text{sub}} \right] dz dx \\
 &= \frac{P_y^2}{4\pi E_S} [3 - 2\gamma - 2 \ln(kR)] + \frac{P_z^2}{4\pi E_S} [5 - 2\gamma - 2 \ln(kR)],
 \end{aligned} \tag{17}$$

and the total potential energy (per unit length) of the whole system including the SiNW and PDMS substrate is

$$\begin{aligned}
 U_{\text{total}} &= U_m + U_b + U_t + U_S \\
 &\quad - \frac{k}{2\pi} \int_0^{2\pi/k} \int_{-R}^R \frac{P_y \cos(kx)}{2R} [w_{\text{sub}} - a \cos(kx)] dz dx, \\
 &\quad - \frac{k}{2\pi} \int_0^{2\pi/k} \int_{-R}^R \frac{P_z \sin(kx)}{2R} [v_{\text{sub}} - b \sin(kx)] dz dx,
 \end{aligned} \tag{18}$$

where the two integral terms in Eq. 18 represent the deformation compatibility between the SiNW and the PDMS substrate. The admissible solution should make the two integral terms equal zero. Substituting Eqs. 6, 9, 14, and 17 into Eq. 18, the total potential energy of the whole system can be written as

$$\begin{aligned}
 U_{\text{total}} &= \frac{EA}{2} \left[\frac{k^2 (a^2 + b^2)}{4} - \varepsilon_{\text{com}} \right]^2 - \frac{[EAak^2 \left(\frac{a^2 k^2}{4} - \varepsilon_{\text{com}} \right) + EIk^4] a}{2} \\
 &\quad - \frac{[EAbk^2 \left(\frac{b^2 k^2}{4} - \varepsilon_{\text{com}} \right) + EIk^4] b}{2} + \frac{EIk^4 (3a^4 k^2 + 2a^2 b^2 k^2 + 3b^4 k^2 + 4a^2 + 4b^2)}{16(a^2 k^2 + a^2 k^4 b^2 + b^2 k^2 + 1)^{\frac{3}{2}}} \\
 &\quad + \frac{AE^2 (a^2 k^2 + b^2 k^2 - 4\varepsilon_{\text{com}})^2 (Aa^2 b^2 k^2 + I_p (\sqrt{(1 + b^2 k^2)(1 + a^2 k^2)} - 1))}{32GI_p \sqrt{(1 + b^2 k^2)(1 + a^2 k^2)}} \\
 &\quad - \frac{[EAak^2 \left(\frac{a^2 k^2}{4} - \varepsilon_{\text{com}} \right) + EIk^4]^2}{4\pi \bar{E}_S} [3 - 2\gamma - 2 \ln(kR)] \\
 &\quad - \frac{[EAbk^2 \left(\frac{b^2 k^2}{4} - \varepsilon_{\text{com}} \right) + EIk^4]^2}{4\pi \bar{E}_S} [5 - 2\gamma - 2 \ln(kR)].
 \end{aligned} \tag{19}$$

As the PDMS substrate is almost volume incompressible, the Poisson's ratio ν_S is set as 0.5. Then, the theoretical solution of the helical buckling spacing and critical

buckling strain can be obtained by minimizing the total potential energy U_{total} with respect to a , b , and k , that is

$$k = \left[\frac{\bar{E}_s}{EI} \frac{2\pi (1 + \eta^2) [1 - \gamma - \ln(kR) + \eta^2(2 - \gamma - \ln(kR))]}{[3 - 2\gamma - 2 \ln(kR) + \eta^2(5 - 2\gamma - 2 \ln(kR))]^2} \right]^{1/4}, \quad (20)$$

and

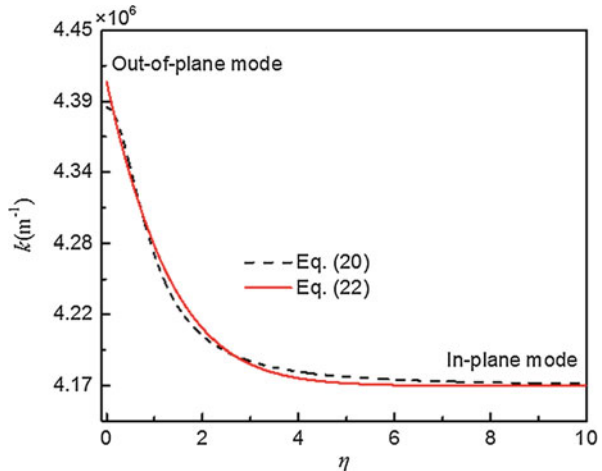
$$\varepsilon_{\text{cr}} = \frac{I}{A} k^2 + \frac{\pi \bar{E}_s (1 + \eta^2)}{A k^2 E [3 - 2\gamma - 2 \ln(kR) + \eta^2 (5 - 2\gamma - 2 \ln(kR))]}, \quad (21)$$

respectively, where $\eta = b/a$ is the ratio between the in-plane displacement and out-of-plane displacement amplitude for the helical buckling of SiNW. If the helical buckling mode degenerates to the out-of-plane ($\eta = 0$) or in-plane ($\eta \rightarrow \infty$) buckling mode, Eq. 20 can degenerate to the buckling wavevector of the out-of-plane or in-plane buckling mode reported in previous work, respectively (Xiao et al. 2008, 2010), which verifies the validity of the theoretical model presented herein. Furthermore, as the right-hand side of Eq. 20 changes very slowly with kR , Eq. 21 can be approximately represented as

$$k = \left(\frac{\bar{E}_s}{EI} \right)^{1/4} \left(\frac{0.4793\eta}{\eta^{2.3} + 23.31} + 0.7399 \right). \quad (22)$$

For the SiNW with Young's modulus 200 GPa and radius of 15 nm and the PDMS substrate with elastic modulus 6 MPa, the relations between the helical buckling wavevector k and η described by Eqs. 20 and 22 are shown in Fig. 3,

Fig. 3 The relations between the wavevector k and η described by Eqs. 20 and 22, respectively.



respectively. It is suggested the relation between k and η can be well described by Eq. 22. The wavevector decreases with the increasing of η and for the perfect helical buckling mode with $\eta = 1$ the wavevector can be approximately given by $k = \frac{3}{4} \left(\frac{\bar{E}_s}{EI} \right)^{1/4}$, which gives the initial helical buckling spacing as

$$h = \frac{2\pi}{k} = \frac{8\pi}{3} (EI/\bar{E}_s)^{1/4}. \tag{23}$$

This relation will be further verified by FEM simulation results in section “FEM Simulations.”

Substituting Eq. 20 into Eq. 21, the critical buckling strain can be rewritten as

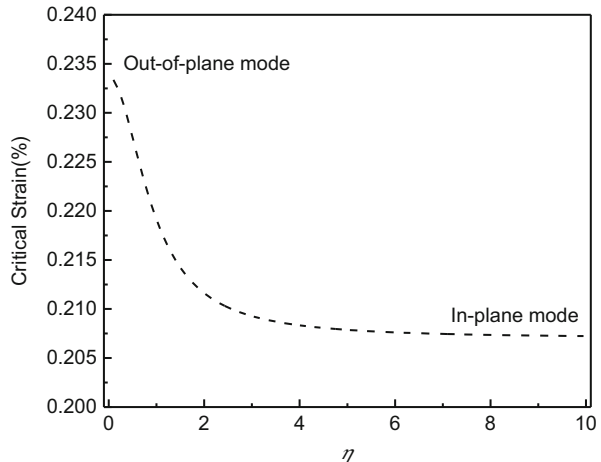
$$\varepsilon_{cr} = \sqrt{\frac{\bar{E}_s}{E}} f(\eta), \tag{24}$$

where $f(\eta)$ is

$$f(\eta) = \sqrt{\frac{(1 + \eta^2)}{8[(1 - \gamma - \ln(kR)) + \eta^2(2 - \gamma - \ln(kR))]} \frac{(5 - 4\gamma - 4 \ln(kR)) + \eta^2(9 - 4\gamma - 4 \ln(kR))}{(3 - 2\gamma - 2 \ln(kR)) + \eta^2(5 - 2\gamma - 2 \ln(kR))}.$$

Interestingly, if we ignore the very slow change of $f(\eta)$ with kR , the critical buckling strain is independent on the radius of the SiNW, R . The relation between the critical buckling strain ε_{cr} and the displacement amplitude ratio η is shown in Fig. 4. Similar to the relation between the buckling wavevector and η , the critical buckling strain

Fig. 4 The relation between the critical buckling strain and the displacement amplitude ratio η described by Eq. 24



decreases with the increasing of η . Besides, the same scaling law $\sqrt{E_S/E}$ of the critical buckling strain has also been found in previous literatures for the in-plane or out-of-plane buckling mode of a nanowire on an elastomeric substrate (Xiao et al. 2008, 2010).

Both experiments and FEM simulations (below) indicate that the number of helical buckling coil in SiNW keeps constant during compression of the PDMS substrate (see Fig. 8). Assuming the initial length of SiNW is l_0 and the number of helical buckling coil is n , the initial helical buckling spacing is $h = l_0(1 - \varepsilon_{cr})/n$ and the post-buckling spacing is $h_p = l_0(1 - \varepsilon_{com})/n$. Thus, the relation between the post-buckling spacing h_p and compressive strain ε_{com} is

$$h_p = h \frac{1 - \varepsilon_{com}}{1 - \varepsilon_{cr}} = \frac{8\pi}{3} \left(\frac{EI}{E_S} \right)^{1/4} \frac{1 - \varepsilon_{com}}{1 - \varepsilon_{cr}}. \quad (25)$$

The post-buckling spacing is determined by the cross section dimension of the nanowire, the Young's modulus of the nanowire and substrate, and the compressive strain, similar to that reported in previous theoretical works (Jiang et al. 2007; Kalita and Somani 2010).

In order to obtain the helical buckling displacement, the total potential energy of the SiNW and PDMS system (Eq. 19) is minimized by conjugate gradient method with respect to a and b . For the effective compressive strain 34%, the out-of-plane and in-plane displacement amplitudes for different diameters of SiNW from 10 nm to 100 nm and different Young's modulus of nanowire from 10 GPa to 250 GPa are shown in Tables 1 and 2, respectively. Here, the elastic modulus of the PDMS substrate is set as 6 MPa. Generally, both of the out-of-plane and in-plane displacement amplitudes a and b increase with the diameter and Young's modulus of the nanowire. Furthermore, the out-of-plane and in-plane displacement amplitudes can be well described by

$$a = \frac{1.259}{k} \sqrt{\varepsilon_{com} - \varepsilon_{cr}} \quad (26)$$

Table 1 The out-of-plane (a) and in-plane (b) displacement amplitudes for different diameters of SiNW

D/nm	10	30	50	70	100
a/nm	59.8	179	299	418	598
b/nm	58.0	174	290	407	580

Table 2 The out-of-plane (a) and in-plane (b) displacement amplitudes for different Young's modulus of nanowire

E/GPa	10	50	100	150	200	250
a/nm	86.7	128	151	167	179	189
b/nm	83	124	147	162	174	184

and

$$b = \frac{1.225}{k} \sqrt{\varepsilon_{\text{com}} - \varepsilon_{\text{cr}}}. \quad (27)$$

The largest discrepancy between the predictions from Eqs. 26 and 27 and the values of a and b obtained by numerically solving Eq. 19 is smaller than 3.5%.

Comparison of the Bending Strain for Different Buckling Modes

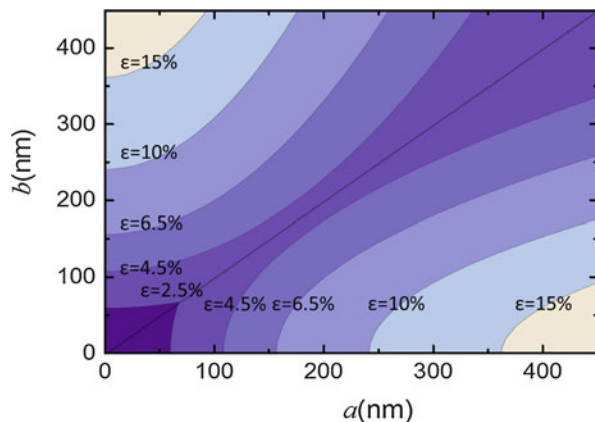
The helical buckling can significantly release the bending strain in SiNW, so that SiNW can sustain larger compression in helical mode than that of the in-plane or out-of-plane modes, enhancing the stretchability (Xu et al. 2011). In this section, the maximum bending strain in SiNW for different buckle modes is analyzed.

As studied in section “Model and Method,” the deflection of SiNW is represented by a general helical form, that is $w = a\cos(kx)$ in y direction and $v = b\sin(kx)$ in z direction. If $a = 0$, it is the in-plane mode; while for $b = 0$, it is the out-of-plane mode. Based on the general helical form, the maximum bending strain in SiNW can be given as (Xu et al. 2011)

$$\varepsilon_{\text{max}} = \begin{cases} aR / \left(b^2 + (h_p/2\pi)^2 \right), & a \geq b \\ bR / \left(a^2 + (h_p/2\pi)^2 \right), & a < b \end{cases}. \quad (28)$$

For the diameter of SiNW $D = 30$ nm at a given post-buckling spacing $h_p = 1.2$ μm (corresponding to an effective compressive strain of $\varepsilon_{\text{com}} = 19.36\%$ for helical buckling). The contour map of ε_{max} for different values of a and b is shown in Fig. 5. The contour map is symmetric along the line of $a = b$ and has the lowest values when $a = b$. Beyond the line $a = b$, ε_{max} increases with the increasing of b (the in-plane displacement amplitude) and decreases with the increasing of a (the out-of-plane displacement amplitude), whereas below the line $a = b$, the trend reverses. Taking the failure strain of SiNWs as 6.5% (Xu et al. 2011; Zhu et al.

Fig. 5 Maximum local strain along the axial direction of SiNW for the SiNW with Young’s modulus 200 GPa, diameter 30 nm, and PDMS with elastic modulus 6 MPa. The line represents the perfect circular buckling mode

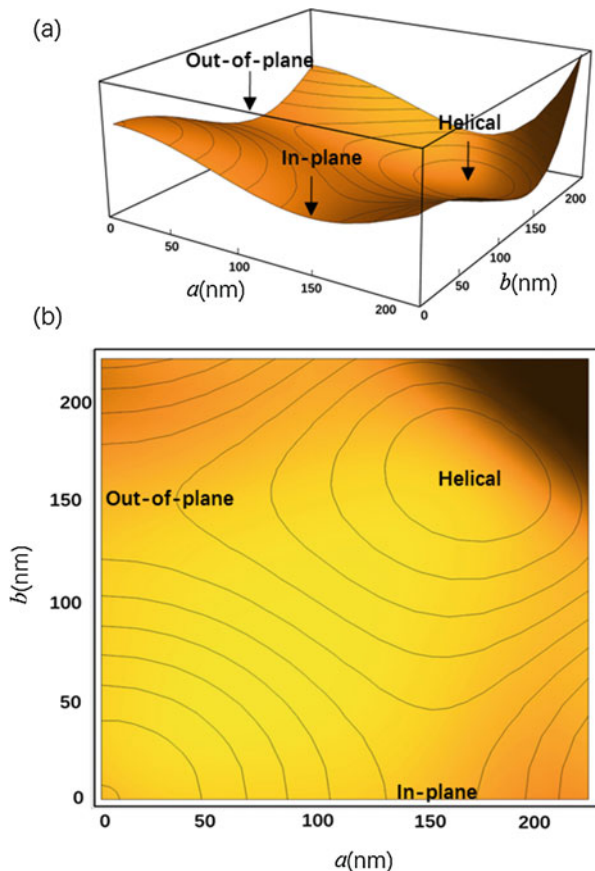


2009), from the contour map we can find that the fracture strain is always larger than ε_{\max} around the line $a = b$, which represents the perfect circular buckling mode, even if some displacement amplitudes (a or b) become large. This means that the stretchability of helical buckled SiNW can be enhanced, which echoes the fact that during experiments the helical buckled SiNW did not fail even upon the failure of PDMS substrate (Xu et al. 2011). The stretchability of helical coil can further benefit by increasing the failure strain of PDMS and bonding strength of the interface between SiNW and PDMS. For $a = b$, another benefit is that the distribution of the bending strain along the axis of the nanowire is uniform which will be discussed in Sect. 4.5.

Energetically Favorable Helical Buckling Mode of SiNWs

Based on the total potential energy of the SiNW and PDMS substrate (Eq. 19), a typical energy landscape for different buckling profiles (i.e., different displacement amplitudes in the in-plane and out-of-plane direction) is given in Fig. 6, in which the Young's modulus of SiNW is set as 200 GPa and the elastic modulus of PDMS

Fig. 6 Energy landscape for different buckling profiles of SiNW on PDMS substrate. (a) Three-dimensional view of the energy landscape and (b) plane view of the energy landscape



is set as 6 MPa, respectively, and the radius of SiNW is 15 nm and the effective compressive strain is 30%. It is shown that both of the in-plane and out-of-plane buckling modes are at the local minimum points of the potential energy landscape. This can explain why in some of the previous experiments the in-plane buckling for a nanowire on an elastomeric substrate has been observed, while in others the out-of-plane buckling has been observed. However, as shown in Fig. 6 the helical buckling has the global minimum potential energy which means the helical buckling is the most favorable buckling mode. Through properly adjusting the loading rate and interface adhesion strength between the nanowire and elastomeric substrate, the helical buckling can be observed in experiments.

FEM Simulations

To verify the theoretical model of helical buckling, FEM simulations using ABAQUS is conducted. For simplicity, we only consider Step 4 in Fig. 1 which is equivalent to applying an effective compressive strain e_{com} to the PDMS substrate. The in-plane dimensions of PDMS substrate is $40 \times 40 \mu\text{m}$ with thickness of $5 \mu\text{m}$, which are much larger than the initial helical buckle spacing and amplitudes of SiNW. SiNW is taken as beams with circular cross section. Besides, the binding between PDMS substrate and SiNW is assumed strong enough without debonding in all simulations. SiNW is consisted of beam elements (B31) with linear elasticity in ABAQUS and for PDMS substrate it is described by linear elasticity with three-dimensional continuum elements (C3D8R), which is consistent with the theoretical model presented in section “[Mechanism of the Transition from In-plane Buckling to Helical Buckling for a Stiff Nanowire on the Surface of an Elastomeric Substrate.](#)” Indeed, the linear elasticity has been widely applied to the PDMS substrate and can give reasonable buckling behaviors of a stiff element on an elastomeric substrate in previous literatures (Mei et al. 2011; Song et al. 2009; Su et al. 2014). We have also conducted the FEM simulation by assuming hyperelastic constitutive relation (Neo-Hooke model) for the PDMS substrate. The helical buckling of the SiNW is also observed with small difference of the buckling spacing and amplitude to that of the PDMS substrate described by linear elasticity. Therefore, in order to directly compare with theoretical results, the linear elasticity has been employed to describe the PDMS substrate. Although the Young’s modulus of SiNWs may be dependent on the diameter of SiNWs (Zhu et al. 2009), such size dependency is ignored in the current study. Different diameter of the SiNW ranging from 10 nm to 100 nm and Young’s modulus ranging from 10 GPa to 250 GPa are also considered in the comprehensive FEM simulations.

Helical Buckling of SiNW on PDMS Substrate

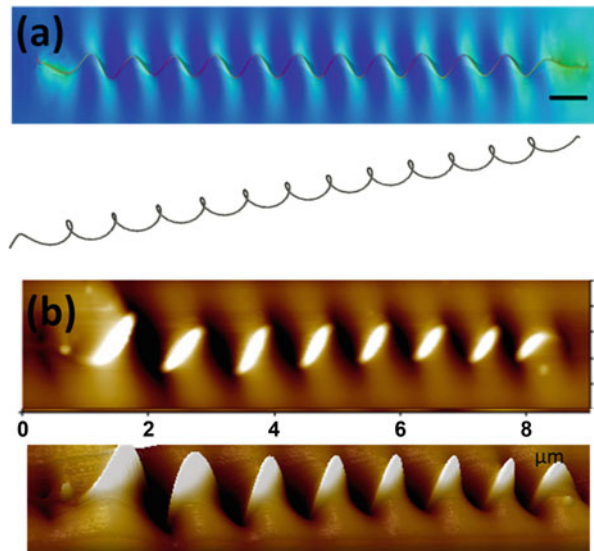
We compare the helical buckle profiles between FEM simulations and experiments to highlight the main geometrical features. The Young’s modulus and diameter of the SiNW used in simulation are 187 GPa for a typical diameter of SiNW 28 nm (Xu et al. 2011; Zhu et al. 2009). As the elastic modulus of PDMS is

varied after being radiated by ultraviolet/ozone (Step 2), in order to fit the initial-buckle spacing measured in experiments ($1.65 \mu\text{m}$) its elastic modulus is set as 3.76 MPa , which is in the range of the elastic modulus of PDMS reported in previous literatures (Ryu et al. 2009; Song et al. 2009; Zhou et al. 2015). The typical simulated buckling configuration of SiNW on PDMS substrate after compression is shown in Fig. 7a, in which the helical coil morphology is easily recognized, and it is noticed that the deformation of the middle of SiNW is uniform. The buckling configuration in parallel experiments is demonstrated through planar and 3D atomic force microscopy (AFM) images of the helical SiNW, as indicated in Fig. 7b.

The evolution of the buckling profile of SiNW on PDMS substrate obtained from FEM simulation is shown in Fig. 8a. As compression proceeds, the displacement amplitudes of the SiNW gradually increase, while the post-buckling spacing gradually decreases, which agrees well with the experimental observations, see Fig. 8b. The two arrows in Fig. 8a are fixed reference points on the SiNW and interestingly; they always correspond to the peaks of the buckling waves during compression. This implies that the peaks always locate at the same points of the SiNW, and the number of coil waves is constant during compression. Further analysis indicates that the relative distribution of the buckling displacement on the SiNW does not change.

The detailed relation between the post-buckling spacing and effective compressive strain for the helical buckling is given in Fig. 8. During helical buckling the number of coils in SiNW is constant, for example, there are always seven coils between the two reference arrows when the compressive strain varies from 5.07% to 26.5% . The relation between the post-buckling spacing and effective compressive strain is given in Fig. 9 for the experiments (red squares), simulation results (black

Fig. 7 The typical configurations of the helical buckling of a SiNW on PDMS substrate. (a) FEM simulations (Top figure: the color contour represents the distribution of the normal strain in x direction ε_{xx} with figure legend presented underneath and the scale bar is $1 \mu\text{m}$. Bottom figure: the three-dimensional configuration of SiNW) and (b) experiments (top figure: planar AFM image. Bottom figure: three-dimensional AFM image. The gray level represents the height of the surface of the PDMS substrate in z direction.)



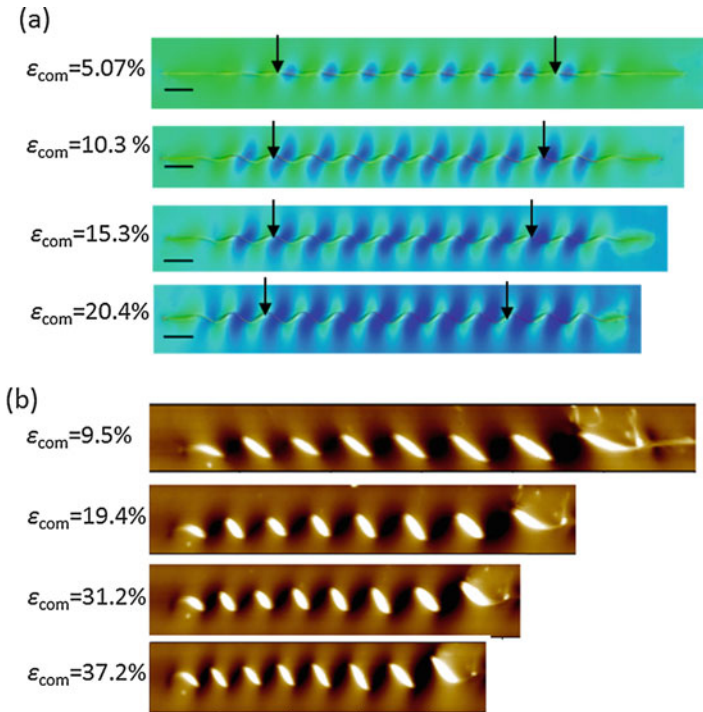


Fig. 8 The snapshots of the typical configuration (top view) for the helical buckling of SiNW on PDMS substrate at different compressive strains. (a) FEM simulations (the color contour represents the distribution of ϵ_{xx}) and (b) experiments. The scale bar is 1 μm

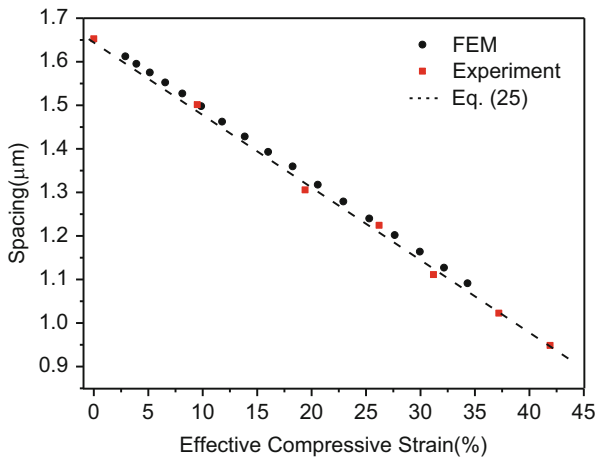


Fig. 9 The post-buckling spacing for helical buckling for different effective compressive strain. The experiments are red squares, simulation results are black circles, and theoretical predictions of Eq. 25 are black dashed line

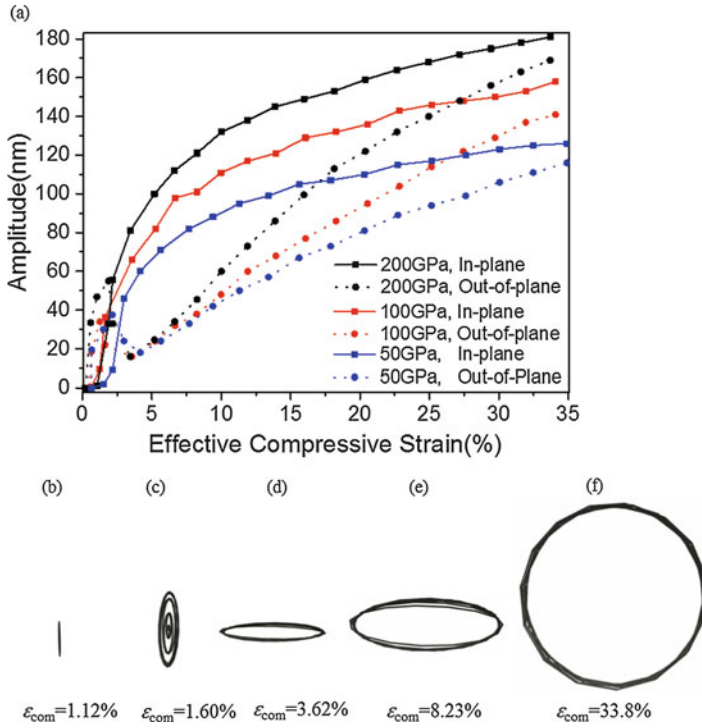


Fig. 10 (a) The FEM simulation results of the in-plane and out-of-plane displacement amplitudes for different effective compressive strain. (b–f) Side views (along the axial direction of SiNW) of the typical buckling profiles of SiNW for several effective compressive strains. The elastic modulus of PDMS is set as 6 MPa, the diameter and Young’s modulus of SiNW are 30 nm and 200 GPa, respectively

circles) and theoretical predictions of Eq. 25 (black dotted line). Nearly perfect accordance with experiments can also verify the effectiveness of both the theoretical model and the FEM results.

Evolution of the Helical Buckling Profile During Compression

FEM simulation results show that the helical buckling profile of SiNW actually changes from a vertical ellipse to a lateral ellipse, and then approaches to a circle when the PDMS substrate is being gradually compressed. Figure 10a shows the amplitude evolution and Fig. 10b–f shows the changes of the SiNW side-view profiles (along the axial direction) with the effective compressive strain. As the effective compressive strain increases exceeding the initial buckling strain, first the out-of-plane displacement increases quickly and the in-plane displacement is almost zero, see Fig. 10a, b. Then, the in-plane displacement becomes more prominent in contrast to slightly decreasing of the out-of-plane displacement as the effective strain gets larger. With further compression, both the in-plane and

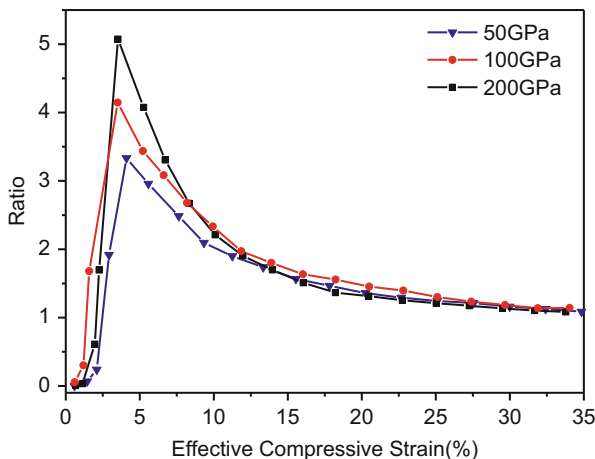


Fig. 11 The FEM simulation results of the ratio of in-plane displacement amplitude to out-of-plane displacement amplitude for the helical buckling of SiNW with different Young's modulus

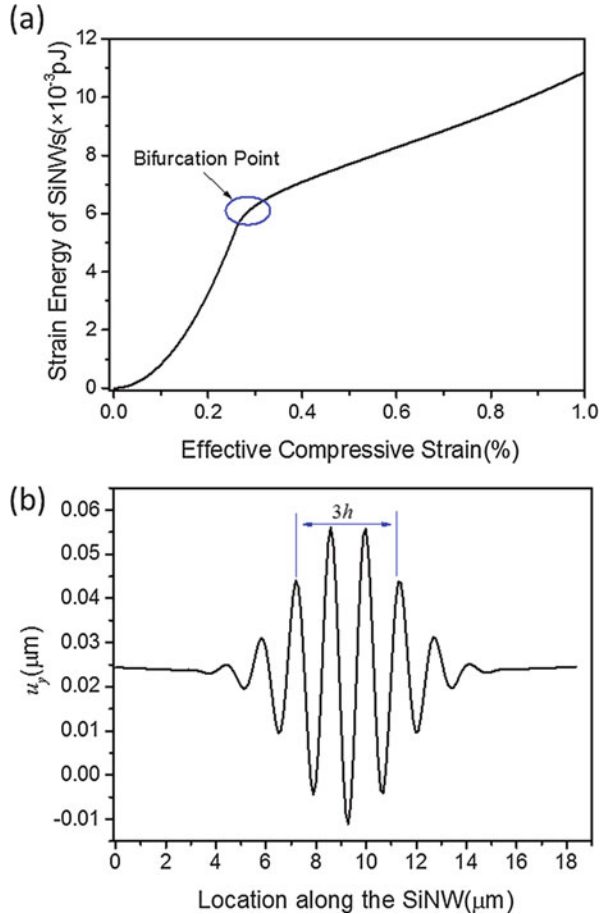
out-of-plane amplitudes increase, and eventually the helical coil profile of SiNW becomes circular to accommodate the large local strain, with the ratio of in-plane amplitudes to out-of-plane amplitudes shown in Fig. 11. In our current FEM simulations, due to the extensively distorted elements in the buckling area, the maximum compressive strain applied to the PDMS substrate is 35% smaller than the experimental value 42%. However, as the buckling profile is gradually approaching to a circle with compression, which is similar to the experiments, we can predict the helical buckling profile is a circle to release the bending strain in SiNW with further increasing the compressive strain.

As shown in Fig. 11, the ratio of in-plane displacement amplitude to out-of-plane displacement amplitude quickly increases with the compressive strain when the compressive strain is larger than the initial buckling strain, and reaches the local acmes at the compressive strain about 4%. Generally, the nanowires with higher Young's modulus have larger peak values of the displacement amplitude ratio. Thereafter, the ratio conspicuously decreases with the magnitude approaching 1. The same trend in the displacement amplitude ratio was observed in experiments (Xu et al. 2011) (note that the minimum effective compressive strain in the experiments was larger than 5%, so only the decreasing trend was able to observe in experiments). Since the displacement amplitude ratio approaching 1 can accommodate larger effective compressive strain (e.g., the circular profile can take the largest strain), the buckled NWs appear to be self-adaptive or "smart" to search for an optimum displacement amplitude ratio.

Helical Buckling Spacing

In order to obtain the initial buckling spacing in FEM simulations, we should first clarify the onset of buckling. Here, the strain energy of SiNW is given in Fig. 12a

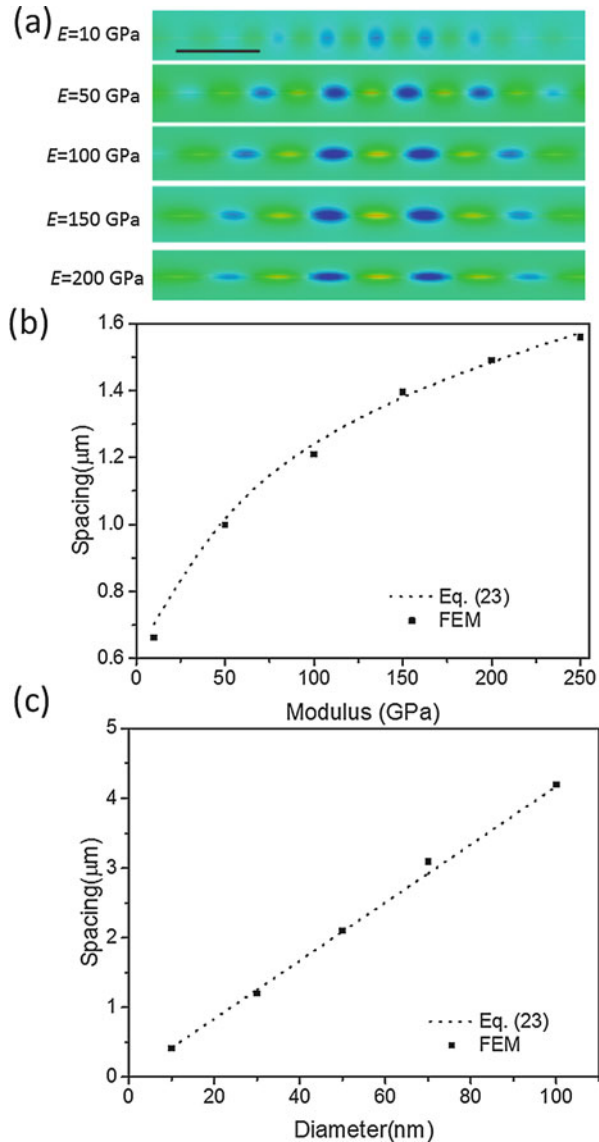
Fig. 12 (a) The strain energy of a SiNW on PDMS substrate during compression in FEM simulation and (b) the out-of-plane displacement of SiNW at the onset of buckling, where the modulus of SiNW and PDMS is 200 GPa and 6 MPa, respectively, and diameter of SiNW is 30 nm



and a clear bifurcation point is found in the strain energy-compressive strain curve. Therefore, the onset of the buckling is defined as the bifurcation point. As the major displacement of SiNW is perpendicular to the surface of PDMS when the compressive strain is smaller than 2%, as shown in Figs. 10 and 11, the out-of-plane displacement of SiNW at the bifurcation point is used to determine the initial buckling spacing, as shown Fig. 12b. For the case shown in Fig. 12b, the out-of-plane displacement v of SiNW is about 50 nm much smaller than the initial buckling spacing, which echoes this point is very close to the initial buckling point. Because of the computational accuracy, it is impossible to determine the exact initial buckling point in FEM simulation. As the influence of the boundary, the buckling amplitude is not uniform. However, the buckle spacing is quite uniform, and the initial buckling spacing h is 1.47 μm , see Fig. 11b.

The contour of axial normal strain (ε_{xx}) on the surface of PDMS corresponding to the bifurcation point in Fig. 12 is given in Fig. 13a. The initial buckling spacing

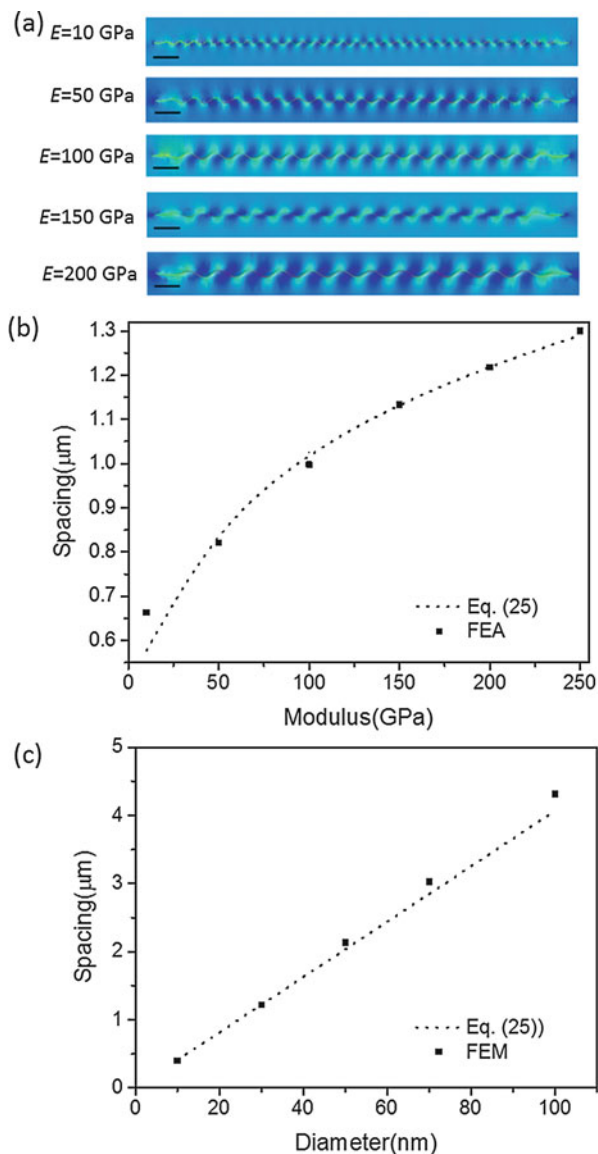
Fig. 13 (a) The axial normal strain ε_{xx} contour (top view) at the onset of buckling shown in Fig. 12. The scale bar is $1\ \mu\text{m}$. The relations of the initial buckling spacing to the Young's modulus of SiNW (b) and diameter of SiNW (c). The dotted line is the theoretical prediction from Eq. 23



increases as the Young's modulus of SiNW. FEM simulation shows that the initial buckling spacing is almost proportional to the $1/4$ power of the Young's modulus of SiNW and is linear to the diameter of SiNW (see Fig. 13b, c), which agrees well with Eq. 23.

After the initial buckling, the deflection of SiNW increases accompanied by the decreasing of the post-buckling spacing as the compression proceeds. Figure 14a shows the axial normal strain contour (ε_{xx}) on the surface of PDMS at the

Fig. 14 (a) The FEM simulation results of the helical buckling configurations at the effective compressive strain of 18% (ϵ_{xx} contour and top view). The scale bar is 1 μm . The relations of the post-buckling spacing to Young's modulus of NW (b) and diameter of SiNW (c). The dotted line is the theoretical prediction from Eq. 25

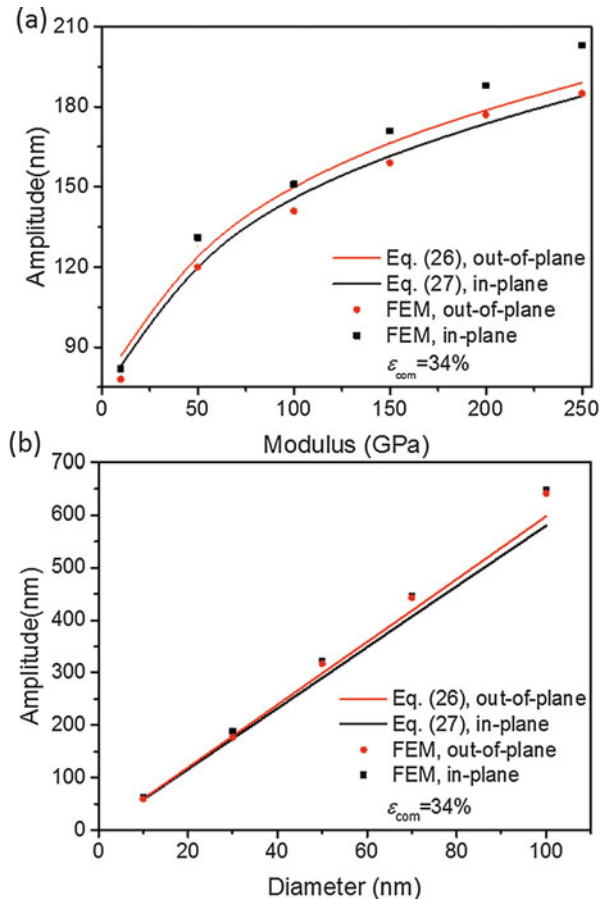


effective compressive strain of 18%. The post-buckling spacing increases with the increasing of SiNW modulus for a given compressive strain, as shown in Fig. 14b, c. Both the theoretical (see Eq. 25) and FEM simulation results indicate that the post-buckle spacing scales with the 1/4 power of the SiNW modulus and increases linearly with SiNW diameter, the same trends as the initial buckling spacing.

Displacement Amplitudes of Helical Buckling

In Fig. 15, the comparisons of the in-plane and out-of-plane displacement amplitudes between the theoretical predictions (Eqs. 26 and 27) and FEM simulations are presented. The FEM simulation results show that the in-plane and out-of-plane displacement amplitudes increase linearly with SiNW diameter and are almost proportional to the 1/4 power of the SiNW modulus, which fit very well to the theoretical predictions. Besides, the displacement amplitudes in y and z directions are almost equal for the effective compressive strain larger than 34%, which means the SiNW buckles into a circular coil form, helping to accommodate larger bending deformation with less maximum local strain. Previous studies indicate that even if the effective compression strain reaches as high as 50%, the maximum local strain in SiNW is less than 3%, considerably lower than the fracture strain of SiNWs (6.5%) (Xu et al. 2011).

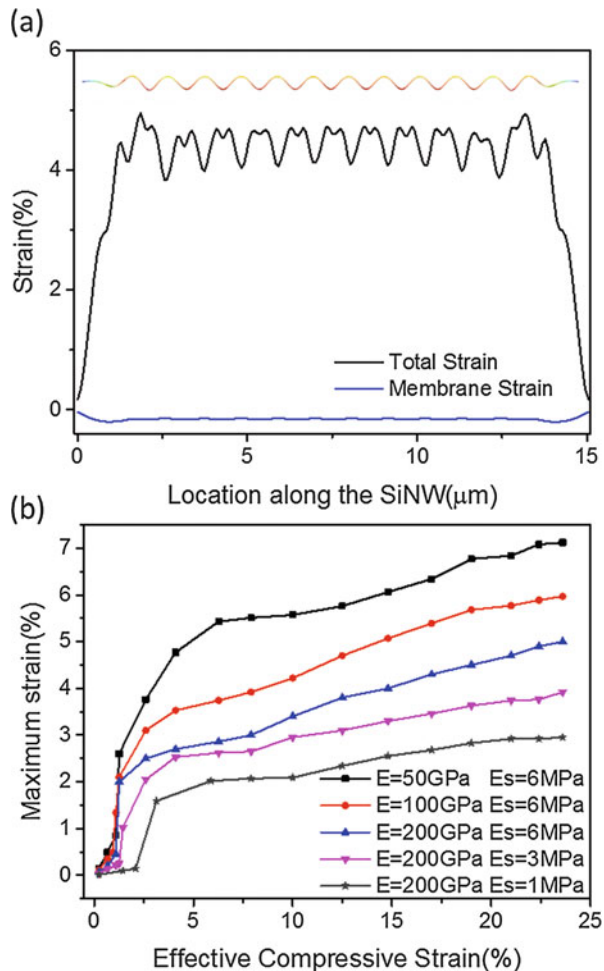
Fig. 15 The in-plane and out-of-plane displacement amplitudes for the helical buckling of SiNW on PDMS substrate versus (a) modulus and (b) diameter of SiNW. The dashed lines are the theoretical predictions from Eqs. 26 and 27



Strain Distribution in SiNW

The advantage of the helical buckling of SiNW can be revealed from the FEM simulation results in Fig. 16a. The membrane strain in SiNW, compared to the total strain (mainly the bending strain), is considerably small and can be regarded as a constant, verifying the assumption in section “Model and Method.” The inset of Fig. 16a shows the distribution of the maximum local strain along the helical buckling configuration. The maximum local strain at the cross section of SiNW is quite uniform and much smaller than the externally imposed compressive strain in the PDMS substrate, which is beneficial to ensure the uniform electronic properties along the SiNW. However, there is still a small fluctuation of the maximum local strain with the fluctuation period equal to the post-buckling spacing, see Fig. 16a. This is because the helical buckling profile of SiNW is not a perfect circle and the

Fig. 16 (a) FEM results of the membrane strain and the maximum local strain at the cross section of SiNW along the axial direction of SiNW for a given compressive strain of 25%. The inset shows the distribution of the maximum local strain in the helical buckling configuration of SiNW. The Young’s modulus and diameter of SiNW are 200 GPa and 30 nm, respectively. (b) The maximum local strain of SiNW versus the effective compressive strain



maximum local strain is the largest at the points of the largest deflection point. The maximum local strain versus the effective compressive strain for different SiNW and PDMS modulus (with constant diameter of SiNW, i.e., 30 nm) is shown in Fig. 16b. At a small effective compressive strain the maximum local strain increases fast, and as the effective compressive strain gets larger, the maximum local strain becomes more stable, which agrees well with the previous studies (Xu et al. 2011).

The nanowire/substrate modulus ratio has a profound impact on the maximum local strain. In Fig. 16b, the maximum local strain for several modulus ratios (E/E_s) are given, and as the modulus ratio decreases, the maximum local strain decreases remarkably. This is because the post-buckling spacing h_p is proportional to the 1/4 power of the nanowire/substrate modulus ratio E/E_s and based on Eq. 28 the larger value of h_p has smaller bending strain. In particular, when the modulus ratio is small, e.g., 1/200,000 (when $E_s = 1$ MPa and $E = 200$ GPa), the maximum local strain is smaller than 3% even at a large effective compressive strain 25%, and such a helical buckled coil profile may accommodate large deformation. This phenomenon is attributed to the deformation homogeneity along the SiNW for helical coil buckling. Note that the fracture strain of SiNW with a diameter of 28 nm is 6.5%, and thus one does not expect fracture to occur even for very large effective compressive strain (such as 50% in previous experiments (Xu et al. 2011)).

Mechanism of the Transition from In-plane Buckling to Helical Buckling for a Stiff Nanowire on the Surface of an Elastomeric Substrate

In most previous works about the buckling behaviors of the nanowire/substrate system, the nanowires were assumed perfectly bonded to the elastomeric substrates throughout deformation, and the effect of the interface debonding was ignored. However, for the buckling of the SiNWs on the PDMS substrate, the UVO treatment time plays an important role in the buckling mode of the SiNWs. The SiNWs may partially debond from the PDMS substrate during compressive buckling, especially with less UVO treatment time. Though the competition between the in-plane buckling mode and out-of-plane buckling mode of SiNW has been studied in recent literature, the selection of the buckling mode is determined by minimum moment of inertia due to the noncircular cross section of the SiNW (Duan et al. 2015). In this section, the buckling behaviors of the SiNWs partially bonded to the PDMS substrate are studied to explore the mechanism of the transition from in-plane buckling to helical buckling.

Model and Method

As shown in Fig. 17, a silicon nanowire (SiNW) is placed on the PDMS substrate and a uniaxial compression is applied on the PDMS substrate in x direction (the axial direction of SiNW); this is equivalent to releasing the prestrain in PDMS substrate.

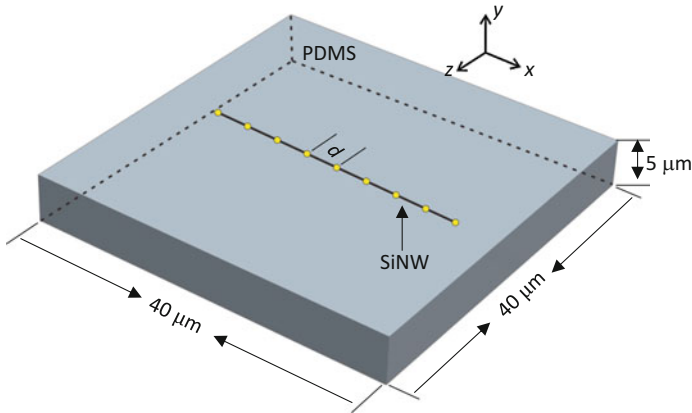


Fig. 17 Model of a SiNW partially bonded to the PDMS substrate. The yellow dots represent the constraint between the SiNW and the PDMS substrate and d is the distance between the adjacent constraint points. A uniaxial compression is applied to the PDMS substrate in x direction

Only the yellow dots in the SiNW are constrained to the PDMS substrate, so as to represent the partial bonding between the SiNW and the PDMS substrate. The constraint between the SiNW and the PDMS is the node-type-constraint, that means the corresponding nodes in the SiNW and the PDMS have the same displacement. Indeed, we have checked different type constraints, e.g., node-type constraint, element-type-constraint, and the influence to the critical buckling strain of SiNW is negligible.

The distance between adjacent constraint points is d , as shown in Fig. 17. As shown in our previous study, the compression buckling mode of a perfectly bonded SiNW ($d/h = 0$) is helical and the buckling spacing is $h = 7.89 R(E_{NW}/\bar{E}_S)^{1/4}$ determined by the radius R , the elastic modulus E_{NW} of the SiNW, and the effective elastic modulus \bar{E}_S of the PDMS substrate. Here, the effective elastic modulus of the substrate is $\bar{E}_S = E_S(1 - \nu_S^2)$, where E_S and ν_S are the elastic modulus and Poisson's ratio of the substrate. Such a helical model pertains when d/h is close to 0, whereas for large d/h , the buckling of the SiNW transfers to the in-plane mode.

Finite element method (FEM) simulations are carried out to study the buckling behaviors of the SiNW on PDMS substrate via the commercial software ABAQUS. The PDMS substrate is simplified as an approximately incompressible isotropic material with Young's modulus $E_S = 3.76$ MPa and Poisson's ratio $\nu_{NW} = 0.475$, and the SiNW is simplified as a beam with circular cross section. The elastic modulus and Poisson's ratio of the SiNW are $E_{NW} = 187$ GPa, $\nu_{NW} = 0.3$ (Zhu et al. 2009). The radius of SiNW is $R = 15$ nm in consistent with parallel experiments. The PDMS substrate and the SiNW are discreted by C3D8R and B31 elements, respectively. In all FEM simulations, the length of the SiNW is $20 \mu\text{m}$, which is much longer than the buckling wavelength of the SiNW studied in this work, and the SiNW lies in the middle of the PDMS surface and far away from the

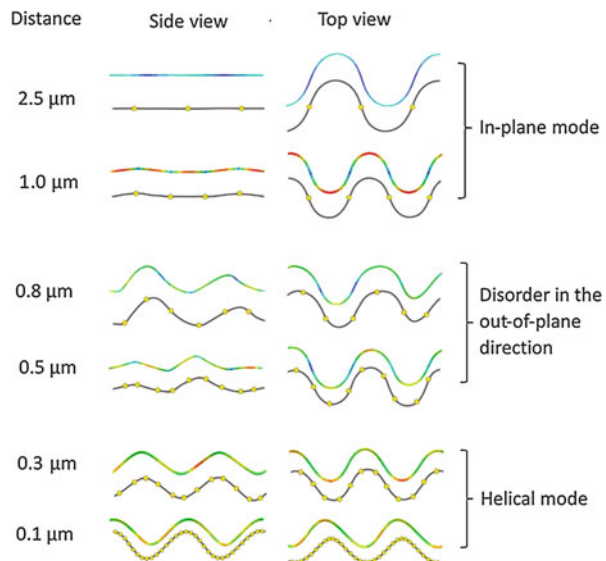
surface edges (which is also consistent with the parallel experiments). Therefore, the boundary effect is regarded small. Although finite boundary is utilized in our FEM simulation, in the middle part of SiNW the buckling configuration exhibits excellent periodicity (see Fig. 23). Mesh convergence is carried out to ensure the reliability of numerical results. Different constraint densities (different d/h) are studied to explore the mechanism of the buckling mode transition.

Results and Discussion

Buckling Modes

The buckling modes of the SiNW for different d are shown in Fig. 18, which can be divided into three distinct regimes as the constraint density increases. When the spacing between adjacent constraint points is large enough, examples such as $d = 2.5 \mu\text{m}$ and $d = 1 \mu\text{m}$, the buckling of the SiNW is in-plane of the PDMS substrate. The in-plane mode is sinusoidal, and the wavelength of the in-plane buckling is $2d$. The maximum strain in the SiNW locates at the maximum curvature point (middle between two constraint points), whereas the strain (as well as curvature) at the constraint point is almost 0. As the constraint density increases, e.g., $d = 0.8 \mu\text{m}$ and $d = 0.5 \mu\text{m}$, the out-of-plane displacement of the SiNW becomes significant. The in-plane displacement is still sinusoid like, but the out-of-plane displacement is disordered. We term this mode as the disordered buckling in the out-of-plane direction. As the constraint density increases further, the configuration of the buckled SiNW becomes the helical coil, as shown for $d = 0.3 \mu\text{m}$ and $d = 0.1 \mu\text{m}$ in Fig. 18. Besides, for the helical buckling the

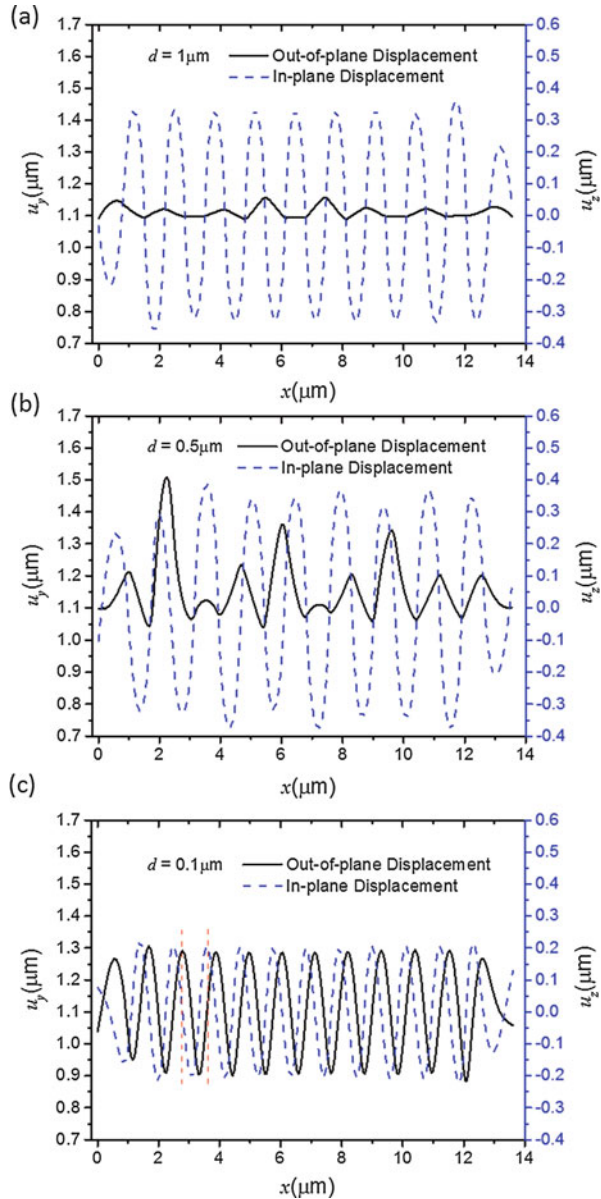
Fig. 18 The buckling modes of the SiNW partially bonded to a PDMS substrate. The yellow dots represent the constraints (bonded sites) between the SiNW and the PDMS substrate. Color contour of the maximum principal strain in the buckled configuration is given



strain distribution in the SiNW is much more uniform than that for the in-plane and disordered modes.

The corresponding in-plane and out-of-plane displacements are shown in Fig. 19 for $d = 1 \mu\text{m}$, $0.5 \mu\text{m}$, and $0.1 \mu\text{m}$, respectively. For $d = 1 \mu\text{m}$, the in-plane displacement is periodic and much larger than the out-of-plane displacement. While

Fig. 19 The in-plane (u_x) and out-of-plane (u_z) displacements of SiNW corresponding to different modes in Fig. 18 ($d = 1 \mu\text{m}$, $0.5 \mu\text{m}$, and $0.1 \mu\text{m}$), respectively



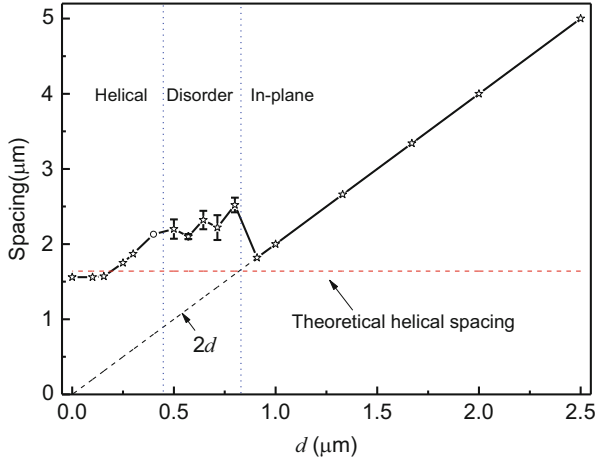


Fig. 20 The buckling wavelength of SiNW for different constraint densities. The red dashed line indicates the theoretical buckling spacing h of the perfectly bonded SiNW and the black dashed line is $2d$. Error bars are given for the out-of-plane disordered buckling mode

for $d = 0.5 \mu\text{m}$, the amplitude of the out-of-plane displacement is of the same order as the in-plane displacement; however, it is fairly irregular (while the in-plane component remains approximately periodic which determines the buckling wavelength). For the helical buckling ($d = 0.1 \mu\text{m}$), the in-plane and out-of-plane displacements are similar. Nevertheless, the phase angle difference of the in-plane and out-of-plane displacements is about 90° , which generates the helical coil form.

The buckle spacing (or buckling wavelength for in-plane buckling) of the SiNW for different constraint densities are shown in Fig. 20, from which the three regimes are obvious. For the in-plane mode, the buckling wavelength is exactly $2d$. While for the disordered mode, the relation between the buckling spacing and the distance d is not monotonous. As the distance d approaches to 0, the profile of the SiNW transfers to the helical buckling, whose wavelength approaches to the theoretical prediction of a perfectly bonded SiNW, $h = 7.89R(E_{\text{NW}}/\bar{E}_S)^{1/4}$. Substituting the parameters used in the present FEM simulation, the theoretical helical buckling spacing is $h = 1.64 \mu\text{m}$. Further parametric studies show that the buckling mode of the SiNW is mainly governed by the ratio d/h . For $d/h > 0.50$, it is the in-plane buckling and for $0.27 < d/h < 0.50$, it is the disordered mode. While, for roughly $d/h < 0.27$, the buckling of the SiNW is predominantly helical.

The critical buckling strain, namely, the compressive strain applied to the PDMS substrate when the SiNW buckling initiates, is shown in Fig. 21. For the in-plane mode, the buckling behavior between adjacent constraint points is analogous to that of a simple supported beam with length d . Based on the Euler buckling theory, the initial buckling strain for a circular beam with radius R is $(\pi R/2d)^2$, which reasonably fits the simulated buckling strain of the in-plane modes. As the constraint density increases, the critical buckling strain gradually diverges from the Euler beam

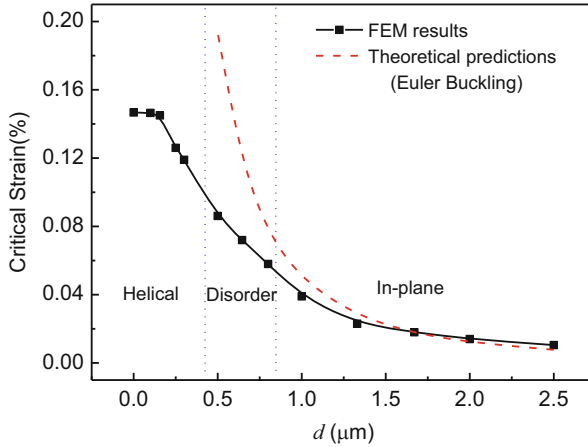


Fig. 21 The critical buckling strain of the partially bonded SiNW with different constraint densities. The red dashed line indicates the Euler beam buckling strain

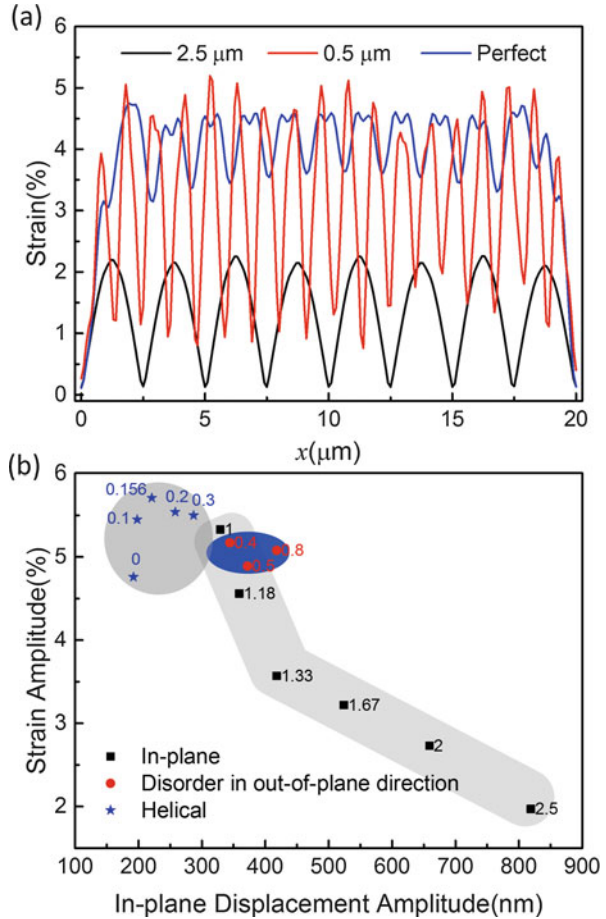
theory, so as to keep a lower strain energy upon buckling. With further increasing the constraint density, the behavior approaches to that of a perfectly bonded SiNW and the initial buckling strain is independent on the distance d .

Strain Distribution in SiNW

The strain distribution in the SiNW is another important factor that influences the performance of the stretchable electronics. For example, fracture may initiate at the largest strain point and the electrical properties of the SiNW are also affected by the strain in it (Peng et al. 2009; Sajjad and Alam 2009). The strain distribution along the axis of the SiNW is shown in Fig. 22a for different constraint densities $d = 2.5 \mu\text{m}$, $0.5 \mu\text{m}$, and $0 \mu\text{m}$. Here, the strain represents the maximum principal in the cross section of the SiNW. The strain distribution for the in-plane mode ($d = 2.5 \mu\text{m}$ for example) fluctuates with the maximum value attained at the middle of two adjacent constraint points, and the minimum value at the constraints. For the disordered mode (e.g., $d = 0.5 \mu\text{m}$), the strain fluctuation is more severe with the maximum exceeding 5% and minimum below 2%. Relatively speaking, the strain distribution in the helical mode is more uniform and oscillates, whose maximum 4.76% is slightly smaller than the disordered mode.

A map of the amplitudes of the maximum strain and the maximum in-plane displacement is given in Fig. 22b; in each case d is denoted. Generally, the strain amplitude increases as the constraint density increases. In contrast, the displacement amplitude decreases. For the in-plane buckling mode, it has smaller strain, but has larger displacement, so that larger space is required and the density of the SiNW is limited. While, for the helical buckling mode it has the smallest displacement and moderate strain amplitude, which might be beneficial for stretchable electronics.

Fig. 22 (a) The strain distributions along SiNW for the perfectly bonded SiNW and partially bonded SiNW ($d = 2.5 \mu\text{m}$, $0.5 \mu\text{m}$) and (b) the maximum strain–maximum displacement relation for different constraint densities

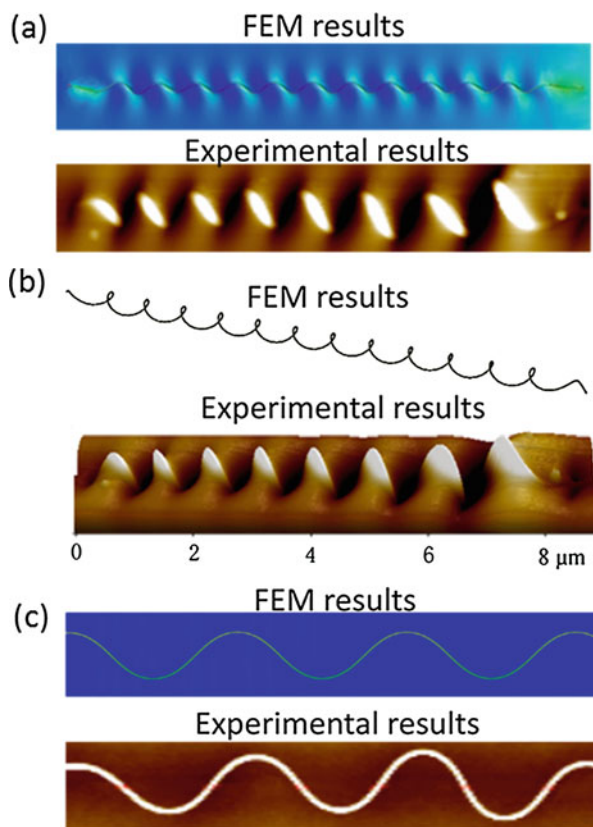


Comparison to Experiments

Parallel experiment is carried out with details outlined in previous literature (Xu et al. 2011). For the SiNW–PDMS system without UVO treatment or less UVO treatment time ($<3 \text{ min}$), the buckling mode of the SiNW is the in-plane which fits the prediction from FEM for a partially bonded SiNW ($d = 1 \mu\text{m}$), as shown in Fig. 23c. When the UVO treatment time is between 5 min and 8 min (when strong adhesion is likely to be present), the buckling of the SiNW is the helical mode which is also consistent with the FEM simulation for a perfectly bonded SiNW, as shown in Fig. 23a, b. However, when the UVO treatment time is longer than 8 min, the mode changes to in-plane again, because the UVO treatment weakens the compliance of the PDMS near the surface which may debond.

For the in-plane mode, the surface of the PDMS substrate remains very smooth, which indicates the sliding between the SiNW and the PDMS substrate. The relations of the in-plane displacement and out-of-plane displacement amplitudes to the effective compressive strain are given in Fig. 24. Both displacement amplitudes

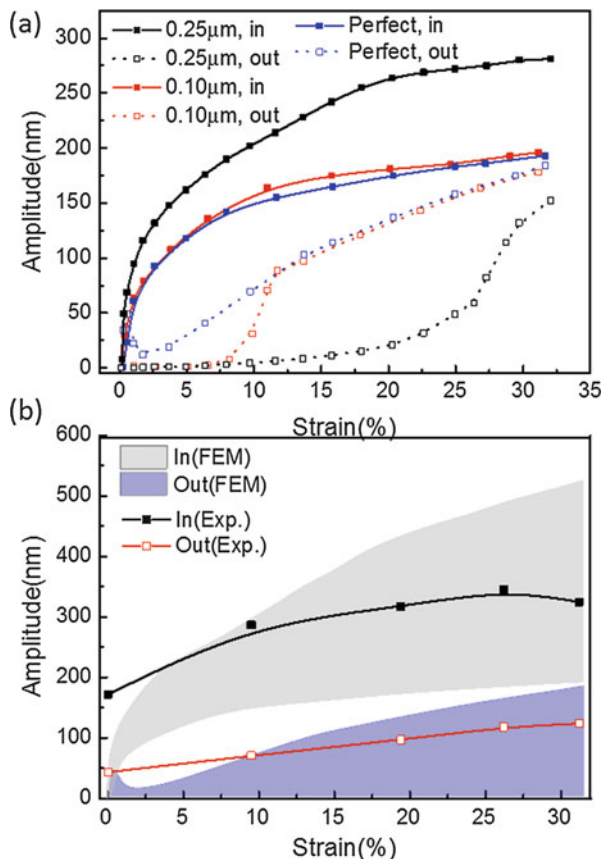
Fig. 23 The typical planar (a) and 3D (b) configuration of the SiNW for the helical buckling and (c) the in-plane buckling mode. The brightness represents the height of the substrate surface



observed in the experiments are within the ranges of the FEM results for the helical buckling ($d = 0.25 \mu\text{m}$, $0.1 \mu\text{m}$ and perfectly bonded SiNW, $d = 0$), as shown in Fig. 24b. This indicates that the partial debonding of the SiNW is the major reason for the transition from the in-plane buckling to the helical buckling. Due to the absent van der Waals force in our model, the present study cannot yet precisely reproduce every feature in experiment; nevertheless, the mechanism of transition of distinctive buckling modes is elucidated.

In order to explore the underlying mechanism of the buckling mode transition for different constraint density, the strain energy of the three different buckling modes (in-plane, disorder, and helical) for three different constraint distances, $2.5 \mu\text{m}$, $0.5 \mu\text{m}$, and $0.1 \mu\text{m}$, is shown in Fig. 25. Here, in order to obtain the in-plane buckling mode for the constraint distance $0.5 \mu\text{m}$ and $0.1 \mu\text{m}$, the out-of-plane displacement of the SiNW is constrained. As shown in Fig. 25, for the constraint distance $0.5 \mu\text{m}$, the strain energy of the in-plane buckling mode is larger than that of the out-of-plane buckling mode. Besides, the strain energy of the out-of-plane disordered buckling mode for the constraint distance $0.5 \mu\text{m}$ is smaller than that of the helical buckling mode for the constraint distance $0.1 \mu\text{m}$. While, for the constraint distance $0.1 \mu\text{m}$, the strain energy of the helical buckling mode is smaller

Fig. 24 (a) The relations of the in-plane and out-of-plane displacement amplitudes to the effective compressive strain for the helical buckling obtained from FEM simulations. (b) The comparison of the in-plane and the out-of-plane displacement amplitudes between experiments (up various UVO treatment time) and FEM simulations. The shadow areas indicate the ranges obtained from FEM simulations



than that of the in-plane buckling mode. Note that we cannot generate the helical buckling mode for the constraint distance $0.5 \mu\text{m}$ and the out-of-plane disordered buckling mode for the constraint distance $0.1 \mu\text{m}$. Here, we cannot directly compare the strain energy between the out-of-plane disordered mode and the helical buckling mode for the constraint distance $0.5 \mu\text{m}$ and $0.1 \mu\text{m}$. Nevertheless, it can be inferred that the out-of-plane disordered buckling mode is energetically more favorable than the in-plane buckling mode for $0.27 < d/h < 0.5$ and the helical buckling mode energy is more favorable for $d/h < 0.27$. In our future work, the underlying mechanism of the buckling mode transition for a nanowire on the elastomeric substrate will be systematically studied.

It should be noted that our FEM model is a simplification to the buckling process. In the experiments, the debonding of the SiNW is accompanied by the buckling of the SiNW, while in our FEM simulation we assume the debonding occurs prior to the buckling. Moreover, the constraints in experiments may be small segments instead of singular points. In the future work, we will systematically study the debonding characteristics of the SiNW and its influence to the buckling

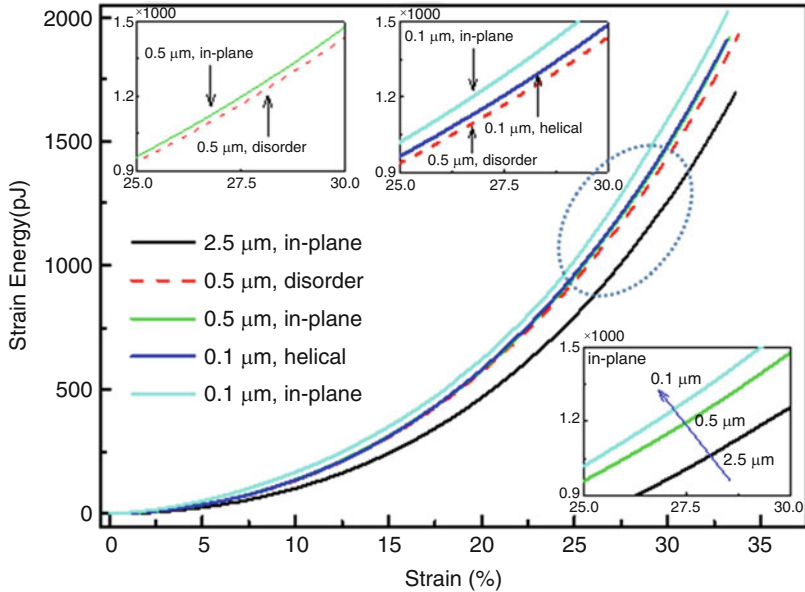


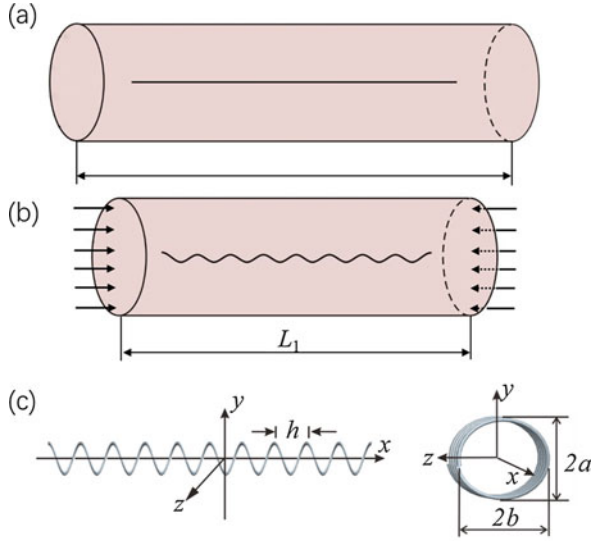
Fig. 25 The strain energy of three buckling modes for the constraint distance 2.5 μm , 0.5 μm , and 0.1 μm , respectively. For the constraint distance 0.5 μm , the strain energy of the in-plane buckling mode (green line) is almost overwritten by the strain energy of the out-of-plane disordered buckling mode (blue line)

behaviors of the SiNW. Besides, the mechanical properties of the PDMS, especially near the surface of the substrate, are influenced by the UVO treatment, for example, increasing the stiffness and decreasing the deformability of the PDMS. In the future work, the gradient of the mechanical properties near the surface of the PDMS will be factored into approach.

Helical Buckling of a Nanowire Embedded in a Soft Matrix Under Axial Compression

In previous sections, we have focused on the 3D helical buckling behaviors and the mechanism of the transition from 2D in-plane buckling to 3D helical buckling of a nanowire on the surface of an elastomeric substrate. Actually, similar 2D and 3D buckling modes were also observed in wires embedded in matrix (Brangwynne et al. 2006; Jiang and Zhang 2008; Li 2008; O’Keeffe et al. 2013; Slesarenko and Rudykh 2016; Su et al. 2014; Zhao et al. 2016). Based on the Winkler foundation model Timoshenko and Gere (2009), and Su et al. (2014) developed the 2D buckling theory of a wire embedded in a soft matrix, and concentrated on the critical strain and initial buckling wavelength (or wavevector) of 2D sinusoidal buckled wires (Slesarenko and Rudykh 2016; Su et al. 2014). However, the 3D helical buckling mechanism and post-buckling behaviors of embedded wires have not been explicitly explained.

Fig. 26 Schematic of helical buckling of embedded wire under axial compression



In this section, a 3D helical buckling theory for compressed wires embedded in a soft matrix is established, and the initial and post-buckling response of wires in association with parallel FEM simulations is discussed systematically.

Theoretical Analysis

Consider a wire with radius R embedded in an infinite matrix (see Fig. 26). According to the experiments of Su et al. (2014), the matrix is so soft that delamination between the wire and the matrix is prevented, and thus the delamination and friction between them is not considered in present research; such simplification is also utilized in their work. When the matrix is compressed and the effective compressive strain (ϵ_{com}) exceeds the critical strain (ϵ_{cr}), the wire buckles into a helical configuration.

For a wire buckling in helical mode, the deformation can be decomposed into two sinusoidal deformations in two perpendicular directions (y and z direction) with x direction denoting the axial direction of the wire. The lateral displacements in y and z direction are assumed as $v = a\cos(kx)$ and $w = b\sin(kx)$, respectively, where a and b are the buckling displacement amplitudes, and k is the buckling wavevector, namely, $2\pi/h$ (h being the initial buckling spacing). The strain energy (per unit length) of the wire can be expressed (Eq. 19) as

$$\begin{aligned}
 U_{wire} = & \frac{EA}{2} \left[\frac{k^2(a^2 + b^2)}{4} - \epsilon_{com} \right]^2 + \frac{EI k^4 (3a^4 k^2 + 2a^2 b^2 k^2 + 3b^4 k^2 + 4a^2 + 4b^2)}{16(a^2 k^2 + a^2 k^4 b^2 + a^2 k^2 + 1)^{\frac{3}{2}}} \\
 & + \frac{(1 + \nu) AE (a^2 k^2 + b^2 k^2 - 4\epsilon_{com})^2 (Aa^2 b^2 k^2 + 2I (\sqrt{(1 + b^2 k^2)(1 + a^2 k^2)} - 1))}{32I \sqrt{(1 + b^2 k^2)(1 + a^2 k^2)}}, \tag{29}
 \end{aligned}$$

where E and ν are the Young's modulus and Poisson's ratio of the wire, respectively. A and I are the cross section area and the moment of inertia of the wire, and ε_{com} is the applied effective compressive strain. The three terms in Eq. 29 represent the membrane strain energy, bending strain energy, and torsion energy of the wire, respectively.

The lateral force experienced by the wire (per unit length) can be given as $T = EI d^4 v / dx^4 - EA \varepsilon_m d^2 v / dx^2$, where the v denotes the lateral deflection and ε_m is the membrane strain. ε_m derived from y - and z -direction sinusoidal deformations can be given as $a^2 k^2 / 4 - \varepsilon_{\text{com}}$ and $b^2 k^2 / 4 - \varepsilon_{\text{com}}$, respectively. Thus, the lateral force in y and z directions can be given as $T_y = -P_y \cos(kx)$ with $P_y = -EA a k^2 (a^2 k^2 / 4 - \varepsilon_{\text{com}}) - EI a k^4$ and $T_z = -P_z \sin(kx)$ with $P_z = -EA b k^2 (b^2 k^2 / 4 - \varepsilon_{\text{com}}) - EI b k^4$. Inside an infinite matrix, a unit point force in y direction at $(x_0, y_0, 0)$ induces a y -directional displacement at $(x, y, 0)$, given by $\frac{1 + \nu_s}{4\pi E_s \sqrt{(x-x_0)^2 + (y-y_0)^2}}$, where E_s and ν_s are the Young's modulus and Poisson's ratio of the matrix. Thus, the y -directional displacement at $(x, y, 0)$ due to the deformation of the wire can be approximately expressed as (Abramowitz and Stegun 1972; Jiang and Zhang 2008; Li 2008; Timoshenko et al. 1970)

$$v_{\text{sub}} = \int_{-R}^R \int_0^\infty \frac{P_y \cos(kx_0)}{2R} \times \frac{1 + \nu_s}{4\pi E_s \sqrt{(x-x_0)^2 + (y-y_0)^2}} dx_0 dy_0. \quad (30)$$

Considering R is much smaller than the buckle spacing h , the dominant term of Eq. 30 can be rewritten into

$$v_{\text{sub}} = \frac{(1 + \nu_s) P_y \cos(kx)}{4\pi R E_s} \{2R(1 - \gamma) + 2R \ln 2 - (R + y) \ln[k(R + y)] - (R - y) \ln[k(R - y)]\}, \quad (31)$$

where $\gamma \approx 0.577$ is the Euler constant.

Then, the strain energy (per unit length) of the matrix due to the sinusoidal deformation of the wire in y direction is

$$\begin{aligned} U_{\text{sy}} &= \frac{k}{2\pi} \int_{-R}^R \int_0^\infty \frac{P_y \cos(kx)}{2R} \cdot v_{\text{sub}} dx dy \\ &= \frac{(1 + \nu_s) P_y^2}{16\pi E_s} \left[\frac{2 - \nu_s}{1 - \nu_s} - 2\gamma - 2 \ln(kR) \right]. \end{aligned} \quad (32)$$

Similarly, the strain energy (per unit length) of the matrix due to the sinusoidal deformation of the wire in z direction can be approximately given as

$$U_{\text{sz}} = \frac{(1 + \nu_s) P_z^2}{16\pi E_s} \left[\frac{2 - \nu_s}{1 - \nu_s} - 2\gamma - 2 \ln(kR) \right]. \quad (33)$$

Thus, the total potential energy (per unit length) of the wire and matrix can be expressed by

$$\begin{aligned}
 U_{\text{total}} &= U_{\text{wire}} + U_{\text{sy}} + U_{\text{sz}} \\
 &- \frac{k}{2\pi} \int_{-R}^R \int_0^{2\pi/k} \frac{P_y \cos(kx)}{2R} [v_{\text{sub}} - a \cos(kx)] dx dy \\
 &- \frac{k}{2\pi} \int_{-R}^R \int_0^{2\pi/k} \frac{P_z \sin(kx)}{2R} [w_{\text{sub}} - b \sin(kx)] dx dy.
 \end{aligned} \tag{34}$$

Here, the two integration terms represent the deformation compatibility between the wire and matrix.

By substituting Eqs. 29, 32, and 33 into Eq. 34, and integrating the last equation, we obtain

$$\begin{aligned}
 U_{\text{total}} &= \frac{EA}{2} \left[\frac{k^2 (a^2 + b^2)}{4} - \varepsilon_{\text{com}} \right]^2 - \frac{[EAak^2 \left(\frac{a^2 k^2}{4} - \varepsilon_{\text{com}} \right) + EIk^4] a}{2} \\
 &- \frac{[EAbk^2 \left(\frac{b^2 k^2}{4} - \varepsilon_{\text{com}} \right) + EIk^4] b}{2} \\
 &+ \frac{(1 + \nu) AE (a^2 k^2 + b^2 k^2 - 4\varepsilon_{\text{com}})^2 (Aa^2 b^2 k^2 + 2I (\sqrt{(1 + b^2 k^2)(1 + a^2 k^2)} - 1))}{32I \sqrt{(1 + b^2 k^2)(1 + a^2 k^2)}} \\
 &- \frac{(1 + \nu) [EAak^2 \left(\frac{a^2 k^2}{4} - \varepsilon_{\text{com}} \right) + EIk^4]^2}{16\pi E_s} [3 - 2\gamma - 2 \ln(kR)] \\
 &- \frac{(1 + \nu) [EAbk^2 \left(\frac{b^2 k^2}{4} - \varepsilon_{\text{com}} \right) + EIk^4]^2}{16\pi E_s} [3 - 2\gamma - 2 \ln(kR)] \\
 &+ \frac{EIk^4 (3a^4 k^2 + 2a^2 b^2 k^2 + 3b^4 k^2 + 4a^2 + 4b^2)}{16(a^2 k^2 + a^2 k^4 b^2 + a^2 k^2 + 1)^{\frac{3}{2}}}
 \end{aligned} \tag{35}$$

Minimizing the potential energy gives the buckling vector as

$$k = \left(\frac{E_s}{EI} \right)^{1/4} \left[\frac{16\pi (1 - \gamma - \ln(kR))}{3(3 - 2\gamma - 2 \ln(kR))^2} \right]^{1/4}. \tag{36}$$

Generally, kR is much smaller than 1, and thus Eq. 36 can be approximated into $k = 1.02 \left(\frac{E_s}{EI} \right)^{1/4}$, and thus the initial buckle spacing is $h = 6.160 \left(\frac{EI}{E_s} \right)^{1/4}$. In particular, for a wire with circular cross section, the initial buckle spacing can also

Table 3 Buckling amplitudes for wires with different Young's modulus ($R = 15$ nm)

E/GPa	10	50	100	150	200	250
a/nm	67.9	101	119	131	141	149
b/nm	67.9	101	119	131	141	149

Table 4 Buckling amplitudes for wires with radius ($E = 200$ GPa)

R/nm	5	10	15	20	25
a/nm	47.0	94.0	141	188	235
b/nm	47.0	94.0	141	188	235

be written into $h = 5.80R\left(\frac{E}{E_s}\right)^{1/4}$. And the post-buckling spacing can be expressed as (Chen et al. 2016; Jiang et al. 2007; Xu et al. 2011):

$$h = 5.80R\left(\frac{E}{E_s}\right)^{1/4} (1 - \varepsilon_{\text{com}}), \quad (37)$$

which shares the similar scaling law $R\left(\frac{E}{E_s}\right)^{1/4}$ as the helical buckling on the surface of substrates, with the prefactor a quarter smaller than the latter.

Substituting Eq. 36 back into Eq. 35, and minimizing with respect to the displacement amplitude a and b , we can numerically obtain these two undetermined profile parameters. The typical amplitudes of wires with different Young's modulus and radius are shown in Tables 3 and 4 (where the effective compressive strain is 35%). The buckling amplitude a and b are equal due to the isotropy of the matrix. Thereafter, we do not differentiate the amplitudes in y and z direction unless specified. By fitting of the theoretical results, the amplitude can be expressed as $a = \frac{1.276}{k}\sqrt{\varepsilon_{\text{com}} - \varepsilon_{\text{cr}}}$ where ε_{cr} is the critical buckling strain. In the present research, ε_{cr} is very small (0.34% for a 15-nm-radius wire of modulus 200 GPa embedded in a matrix of 6 MPa) and thus negligible compared with the effective compressive strain ε_{com} in the post-buckling phase. Therefore, the buckling amplitude can also be simplified as

$$a = \frac{1.276}{k}\sqrt{\varepsilon_{\text{com}}}. \quad (38)$$

Results and Discussion

To verify the theoretical solution, FEM simulations based on commercial software ABAQUS are conducted. The matrix is simplified as a cylinder, which is much larger than the buckling amplitude and spacing. The wire is regarded as a beam with the radius 15 nm along the axial of the cylindrical matrix. Both ends of the cylindrical matrix are slowly compressed toward each other. The Young's modulus

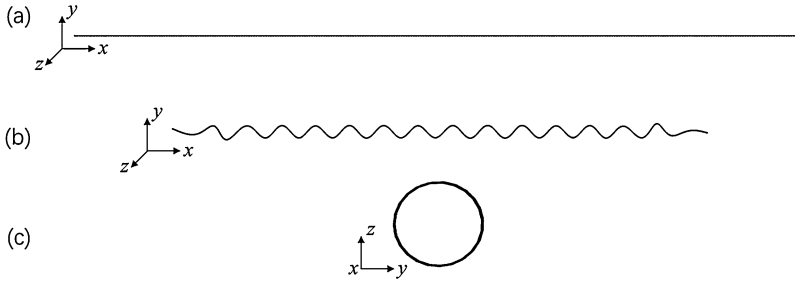
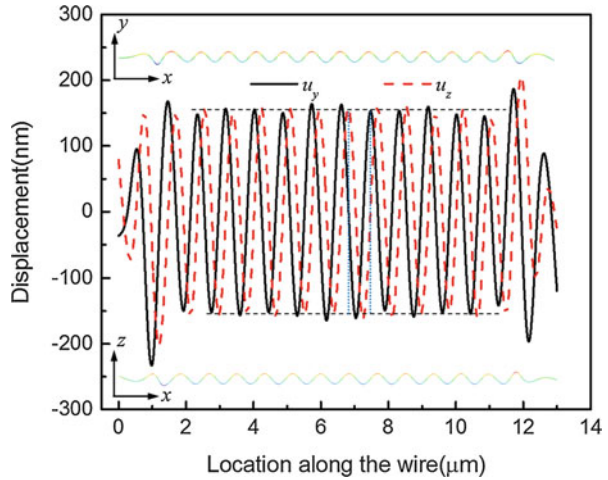


Fig. 27 Typical configurations of the wire before (a) and after (b and c) buckling obtained from FEM simulation

Fig. 28 Typical lateral buckling displacements of a wire obtained from FEM simulation



of the wire is 200 GPa, and 6 MPa for the matrix. Mesh convergence is carried out to ensure the reliability of numerical results. Typical buckling configurations of buckled wires are shown in Fig. 27.

Shown in Fig. 28 are the FEM results of typical buckling displacements of the wire in y and z directions. In both x - y and x - z planes, the projections of buckled wires exhibit the sinusoidal curves. View from the axial direction reveals that the profile of the wire is almost perfect circular, since the buckling amplitudes in x - y and x - z planes are almost equal (also see Fig. 27c). Besides, it should also be noticed that the angle phases in the two planes have a difference of $\pi/2$, which further substantiates the helical configuration of the wire.

The consistency between FEM simulation and theoretical result is shown in Fig. 29, in which the effective compressive strain is 35%, and the theoretical results are Eqs. 37 and 38 for the buckling spacing and amplitude, respectively. The spacing is proportional to wire radius, and the amplitude scales with $1/4$ power of the Young's modulus of the wire, which agrees with previous studies (Chen et al. 2016; Jiang and Zhang 2008; Xiao et al. 2008, 2010; Zhao et al. 2016).

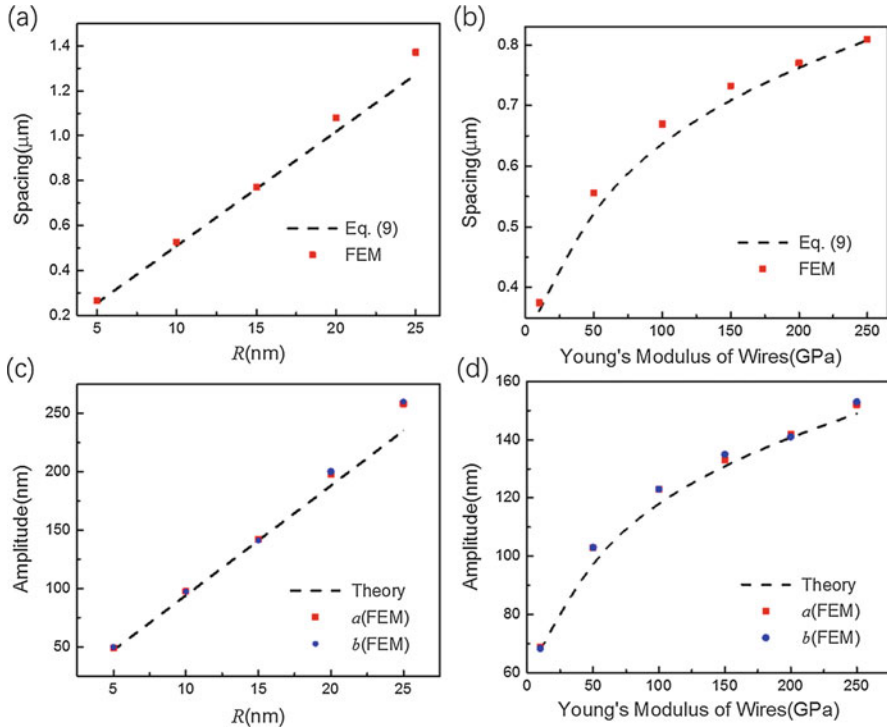
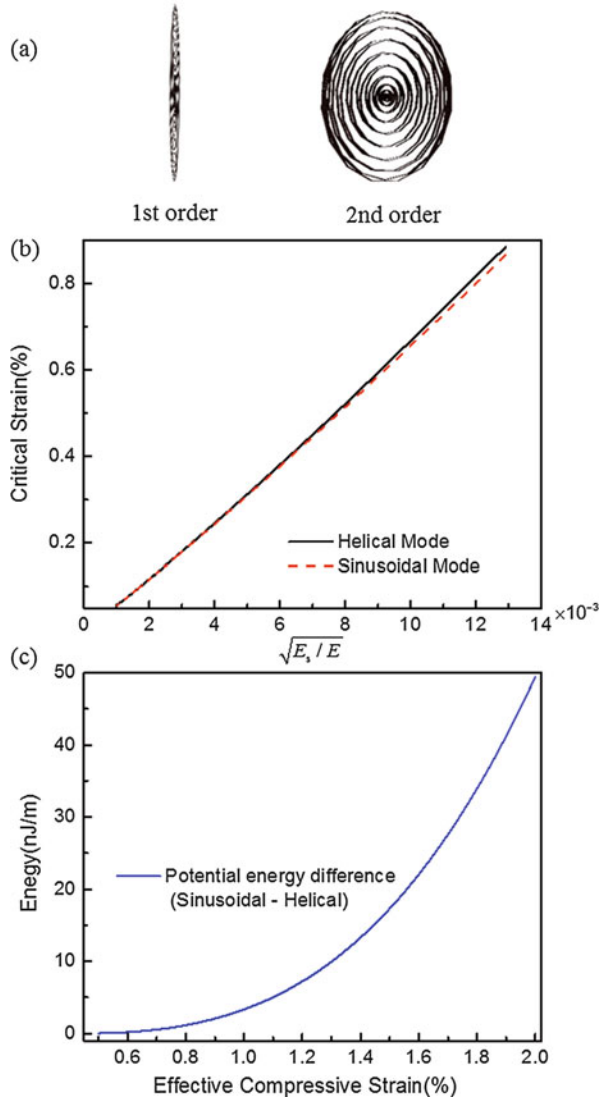


Fig. 29 Post-buckling spacing (**a** and **b**) and amplitude (**c** and **d**) versus the radius (**a** and **c**) and Young's modulus (**b** and **d**) of wires. In (**a**) and (**c**), the Young's modulus of the wire is 200 GPa, and in (**b**) and (**d**) the radius of the wire is 15 nm

The post-buckling behaviors, mainly the buckling spacing and amplitudes, are shown in Fig. 29, which again validates the theoretical model with respect to FEM simulation. The buckling spacing decreases almost linearly with the compressive strain, while the evolution of buckling amplitude is slightly more complex. Once the compressive strain exceeds the critical strain, the sinusoidal buckles initiate mainly in one direction (i.e., almost in a 2D profile); however, the deflection in the other direction quickly catches up and transits the configuration from 2D sinusoidal to 3D helical mode. As shown in Fig. 30, the modal analysis conducted via FEM indicates that the 2D buckling is the first order mode while the 3D helical form is the second order. Thus, the 2D sinusoidal appears first and exhibits lower critical strain, which is also confirmed by the present theory (see Fig. 30b); thereafter, the buckled wire turns into the helical mode with less strain energy. Besides, it should be noted that the critical strain is almost linear with respect to the root of the ratio $(E_s/E)^{1/2}$, which agrees with previous studies (Chen et al. 2016; Su et al. 2014; Xiao et al. 2008, 2010).

As suggested by the FEM results, the amplitude increases with the compressive strain, in accordance with the theoretical prediction. It is noteworthy that even at

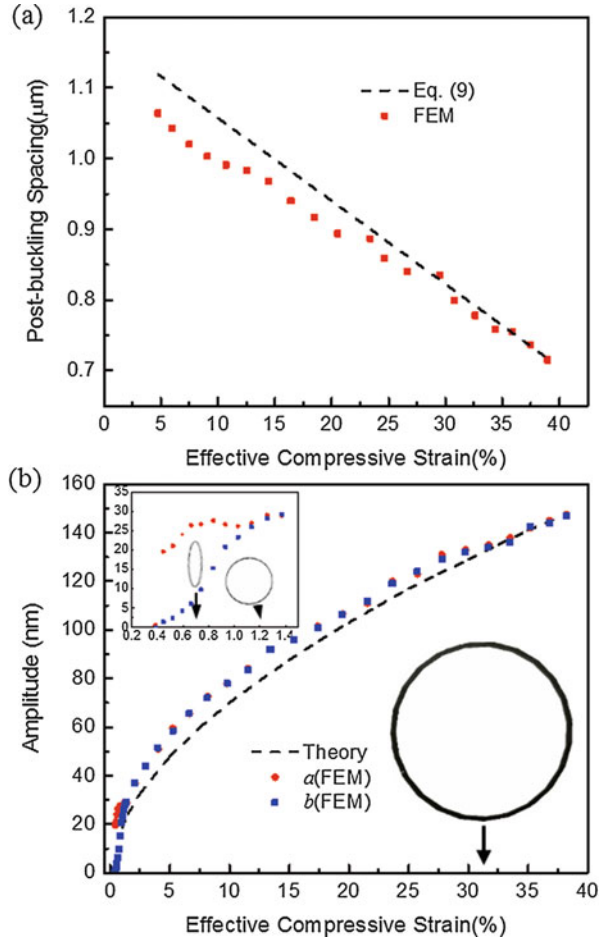
Fig. 30 (a) Modal of the first two orders of the wire embedded in matrix (FEM). (b) Critical strains for the helical buckling mode and sinusoidal mode ($R = 15$ nm). (c) Difference of potential energy (per length) of matrix and wires in the helical mode and sinusoidal mode ($E = 200$ GPa, $E_s = 6$ MPa, $R = 15$ nm)



very small strain (less than 1.5% in the inset of Fig. 31b) the axial view of the buckling configuration turns into almost perfectly circular. By comparison, if the wire sits on top of the surface of a semi-infinite substrate, the axial profile change from an oval toward a circle (Xu et al. 2011) could require a much higher effective compressive strain (over 30%).

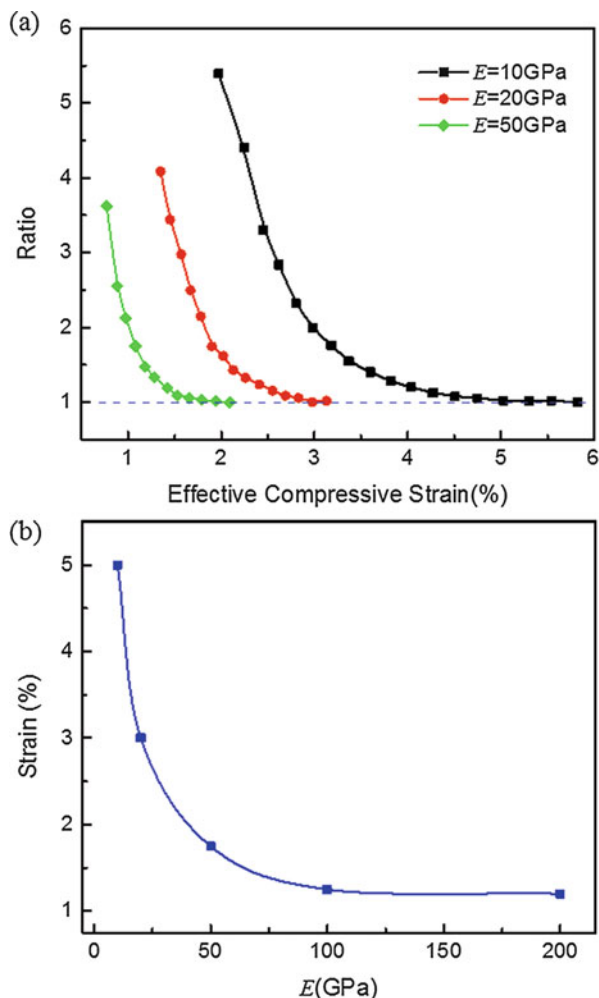
Transition from in-plane buckling mode to helical buckling mode can also be represented by the ratio of in-plane displacement amplitude to out-of-plane

Fig. 31 Theoretical and FEM evolutions of post-buckling behaviors ($E = 200$ GPa, $E_s = 6$ MPa, $R = 15$ nm)



displacement amplitude of wire, as shown in Fig. 32a. Here, the in-plane direction is referred as the initial deformation direction of the wire and the out-of-plane direction is perpendicular to the in-plane direction and the axis of the wire. It is suggested that, when the effective compressive strain reaches a critical value, the buckling of the wire initiates and the ratio is maximum at the beginning (i.e., the in-plane mode); with further compression, the ratio decreases dramatically and approaches 1 (the blue dashed line, i.e., the helical mode). More interestingly, for wires with lower Young's modulus, the effective compressive strain necessary for the transition from 2D to 3D helical mode is higher (see Fig. 32b). Nevertheless, when the Young's modulus is high enough, the transition strain tends to be 1.25%, namely, as the effective compressive strain exceeds 1.25%, wires will buckle into the 3D helical mode.

Fig. 32 (a) The FEM evolution of the ratio of in-plane to out-of-plane displacement amplitude of wire with different Young's modulus and (b) the strain necessary for transition from in-plane mode to helical mode ($E_s = 6$ MPa, $R = 15$ nm)



Conclusion

In summary, we have introduced systematic studies of a special buckling mode – three-dimensional helical coil buckling of SiNW on the UVO-treated PDMS substrate. A continuum mechanics theory based on the minimization of the total potential energy of the SiNW and PDMS substrate is established and the impact of SiNW properties and the geometric dimension on the buckling characteristics, such as the buckling spacing, the in-plane and out-of-plane displacement amplitudes, etc., has been revealed, which agrees quite well with the FEM simulation results. Both the theory and FEM simulation suggest that the buckling spacing and displacement amplitudes increase linearly with the SiNW diameter and are almost proportional to the $1/4$ power of the ratio between the SiNW modulus and the substrate modulus.

In addition, the FEM simulation results indicate that the amplitude ratio changes with the effective compressive strain, and when the effective compressive strain is large enough, the buckling shape is almost circular, namely, the amplitude ratio approaching 1.

The partial detachment of SiNW from the PDMS substrate is mainly responsible for the transition of the buckling modes, which can be dictated by the constraint ratio d/h . For $d/h > 0.50$, the buckling of the SiNW is the in-plane which can be described by the Euler beam buckling theory with the buckling wavelength $\lambda = 2d$. For $0.27 < d/h < 0.50$, the buckles are disordered in the out-of-plane direction and the critical buckling strain gradually diverges from the Euler beam theory. Whereas, for $d/h < 0.27$, the buckling is helical and the buckling spacing gradually approaches to the theoretical value of a perfectly bonded SiNW.

Meanwhile a theoretical analysis of three-dimensional helical buckling of wires embedded in matrix is also provided and the buckling spacing and amplitudes are deduced, which are further verified by parallel FEM simulations. It is suggested that the buckled profile is almost perfectly circular in the axial direction; with increasing compression, the buckling spacing decreases almost linearly, while the amplitude scales with the $1/2$ power of the compressive strain. Besides the transition strain from 2D mode to 3D helical mode decreases with the Young's modulus of the wire and approaches to 1.25% when the modulus is high enough. This study may shed some lights on the buckling behaviors of wires embedded in matrix and provide some useful instructions of manufacturing complex structures.

References

- M. Abramowitz, I.A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (National Bureau of Standards Applied Mathematics Series 55. Tenth Printing. Engineering: 1076, Washington, DC, 1972)
- B. Audoly, A. Boudaoud, Buckling of a stiff film bound to a compliant substrate – part I. *J. Mech. Phys. Solids* **56**(7), 2401–2421 (2008)
- C.P. Brangwynne, F.C. MacKintosh, S. Kumar, N.A. Geisse, J. Talbot, L. Mahadevan, K.K. Parker, D.E. Ingber, D.A. Weitz, Microtubules can bear enhanced compressive loads in living cells because of lateral reinforcement. *J. Cell Biol.* **173**(5), 733–741 (2006)
- B. Charlot, W. Sun, K. Yamashita, H. Fujita, H. Toshiyoshi, In-plane bistable nanowire for memory devices, in *Symposium on Design, Test, Integration and Packaging of MEMS/MOEMS, 2008. MEMS/MOEMS 2008* (IEEE, 2008). <http://ieeexplore.ieee.org/document/4752995/>
- X. Chen, J. Yin, Buckling patterns of thin films on curved compliant substrates with applications to morphogenesis and three-dimensional micro-fabrication. *Soft Matter* **6**(22), 5667–5680 (2010).
- Z. Chen, B. Cotterell, W. Wang, The fracture of brittle thin films on compliant substrates in flexible displays. *Eng. Fract. Mech.* **69**(5), 597–603 (2002)
- Y. Chen, Y. Liu, Y. Yan, Y. Zhu, X. Chen, Helical coil buckling mechanism for a stiff nanowire on an elastomeric substrate. *J. Mech. Phys. Solids* **95**, 25–43 (2016)
- G. Crawford, *Flexible Flat Panel Displays* (Wiley, Chichester, 2005)
- Y. Duan, Y. Huang, Z. Yin, Competing buckling of micro/nanowires on compliant substrates. *J. Phys. D: Appl. Phys.* **48**(4), 045302 (2015)
- J.W. Durham 3rd, Y. Zhu, Fabrication of functional nanowire devices on unconventional substrates using strain-release assembly. *ACS Appl. Mater. Interfaces* **5**(2), 256–261 (2013)

- K. Efimenko, W.E. Wallace, J. Genzer, Surface modification of Sylgard-184 poly(dimethyl siloxane) networks by ultraviolet and ultraviolet/ozone treatment. *J. Colloid Interface Sci.* **254**(2), 306–315 (2002)
- A. Goriely, R. Vandiver, M. Destrade, Nonlinear Euler buckling. *Proc. R. Soc. Lond. A Math. Phys. Eng. Sci.* **464**(2099), 3003–3019 (2008)
- H. Jiang, J. Zhang, Mechanics of microtubule buckling supported by cytoplasm. *J. Appl. Mech.* **75**(6), 061019 (2008)
- H. Jiang, D.Y. Khang, J. Song, Y. Sun, Y. Huang, J.A. Rogers, Finite deformation mechanics in buckled thin films on compliant supports. *Proc. Natl. Acad. Sci. U. S. A.* **104**(40), 15607–15612 (2007)
- S.J. Kalita, V. Somani, $\text{Al}_2\text{TiO}_5\text{-Al}_2\text{O}_3\text{-TiO}_2$ nanocomposite: structure, mechanical property and bioactivity studies. *Mater. Res. Bull.* **45**(12), 1803–1810 (2010)
- D.H. Kim, J. Song, W.M. Choi, H.S. Kim, R.H. Kim, Z. Liu, Y.Y. Huang, K.C. Hwang, Y.W. Zhang, J.A. Rogers, Materials and noncoplanar mesh designs for integrated circuits with linear elastic responses to extreme mechanical deformations. *Proc. Natl. Acad. Sci. U. S. A.* **105**(48), 18675–18680 (2008)
- H.C. Ko, M.P. Stoykovich, J. Song, V. Malyarchuk, W.M. Choi, C.J. Yu, J.B. Geddes 3rd, J. Xiao, S. Wang, Y. Huang, J.A. Rogers, A hemispherical electronic eye camera based on compressible silicon optoelectronics. *Nature* **454**(7205), 748–753 (2008)
- S.P. Lacour, J. Jones, S. Wagner, T. Li, Z. Suo, Stretchable interconnects for elastic electronic surfaces. *Proc. IEEE* **93**(8), 1459–1467 (2005)
- T. Li, A mechanics model of microtubule buckling in living cells. *J. Biomech.* **41**(8), 1722–1729 (2008)
- H. Mei, C.M. Landis, R. Huang, Concomitant wrinkling and buckle-delamination of elastic thin films on compliant substrates. *Mech. Mater.* **43**(11), 627–642 (2011)
- S.G. O’Keeffe, D.E. Moulton, S.L. Waters, A. Goriely, Growth-induced axial buckling of a slender elastic filament embedded in an isotropic elastic matrix. *Int. J. Non Linear Mech.* **56**, 94–104 (2013)
- W.A. Oldfather, C.A. Ellis, D.M. Brown, Leonhard Euler’s elastic curves. *Isis* **20**(1), 72–160 (1933)
- X.H. Peng, A. Alizadeh, S.K. Kumar, S.K. Nayak, Ab initio study of size and strain effects on the electronic properties of Si nanowires. *Int. J. Appl. Mech.* **1**(3), 483–499 (2009)
- Q. Qin, Y. Zhu, Static friction between silicon nanowires and elastomeric substrates. *ACS Nano* **5**(9), 7404–7410 (2011)
- J.A. Rogers, T. Someya, Y. Huang, Materials and mechanics for stretchable electronics. *Science* **327**(5973), 1603–1607 (2010)
- S.Y. Ryu, J. Xiao, W.I. Park, K.S. Son, Y.Y. Huang, U. Paik, J.A. Rogers, Lateral buckling mechanics in silicon nanowires on elastomeric substrates. *Nano Lett.* **9**(9), 3214–3219 (2009)
- R.N. Sajjad, K. Alam, Electronic properties of a strained 100 silicon nanowire. *J. Appl. Phys.* **105**(4), 044307 (2009)
- V. Slesarenko, S. Rudykh, Microscopic and macroscopic instabilities in hyperelastic fiber composites. *J. Mech. Phys. Solids* **99**, 471–482 (2016)
- T. Someya, T. Sekitani, S. Iba, Y. Kato, H. Kawaguchi, T. Sakurai, A large-area, flexible pressure sensor matrix with organic field-effect transistors for artificial skin applications. *Proc. Natl. Acad. Sci. U. S. A.* **101**(27), 9966–9970 (2004)
- J. Song, Y. Huang, J. Xiao, S. Wang, K.C. Hwang, H.C. Ko, D.H. Kim, M.P. Stoykovich, J.A. Rogers, Mechanics of noncoplanar mesh design for stretchable electronic circuits. *J. Appl. Phys.* **105**(12), 123516 (2009)
- C.M. Stafford, C. Harrison, K.L. Beers, A. Karim, E.J. Amis, M.R. VanLandingham, H.C. Kim, W. Volksen, R.D. Miller, E.E. Simonyi, A buckling-based metrology for measuring the elastic moduli of polymeric thin films. *Nat. Mater.* **3**(8), 545–550 (2004)
- T. Su, J. Liu, D. Terwagne, P.M. Reis, K. Bertoldi, Buckling of an elastic rod embedded on an elastomeric matrix: planar vs. non-planar configurations. *Soft Matter* **10**(33), 6294–6302 (2014)

- S.P. Timoshenko, J.M. Gere, *Theory of Elastic Stability* (Courier Corporation, North Chelmsford, 2009)
- S.P. Timoshenko, J.N. Goodier, H.N. Abramson, Theory of elasticity (3rd ed.) *J. Appl. Mech.* **37**(3), 888 (1970)
- Y. Wang, J. Song, J. Xiao, Surface effects on in-plane buckling of nanowires on elastomeric substrates. *J. Phys. D: Appl. Phys.* **46**(12), 125309 (2013)
- E.A. Wilder, S. Guo, S. Lin-Gibson, M.J. Fasolka, C.M. Stafford, Measuring the modulus of soft polymer networks via a buckling-based metrology. *Macromolecules* **39**(12), 4138–4143 (2006)
- J. Xiao, H. Jiang, D.Y. Khang, J. Wu, Y. Huang, J.A. Rogers, Mechanics of buckled carbon nanotubes on elastomeric substrates. *J. Appl. Phys.* **104**(3), 033543 (2008)
- J. Xiao, S.Y. Ryu, Y. Huang, K.C. Hwang, U. Paik, J.A. Rogers, Mechanics of nanowire/nanotube in-surface buckling on elastomeric substrates. *Nanotechnology* **21**(8), 85708 (2010)
- F. Xu, W. Lu, Y. Zhu, Controlled 3D buckling of silicon nanowires for stretchable electronics. *ACS Nano* **5**(1), 672–678 (2011)
- S. Xu, Z. Yan, K.I. Jang, W. Huang, H.R. Fu, J. Kim, Z. Wei, M. Flavin, J. McCracken, R. Wang, A. Badea, Y. Liu, D.Q. Xiao, G.Y. Zhou, J. Lee, H.U. Chung, H.Y. Cheng, W. Ren, A. Banks, X.L. Li, U. Paik, R.G. Nuzzo, Y.G. Huang, Y.H. Zhang, J.A. Rogers, Assembly of micro/nanomaterials into complex, three-dimensional architectures by compressive buckling. *Science* **347**(6218), 154–159 (2015)
- J. Yin, X. Chen, Buckling of anisotropic films on cylindrical substrates: insights for self-assembly fabrication of 3D helical gears. *J. Phys. D: Appl. Phys.* **43**(11), 115402 (2010)
- E. Zeeman, Euler buckling, in *Structural Stability, the Theory of Catastrophes, and Applications in the Sciences* (Springer, 1976), pp. 373–395. <http://www.springer.com/la/book/9783540077916>
- Y. Zhao, J. Li, Y.P. Cao, X.Q. Feng, Buckling of an elastic fiber with finite length in a soft matrix. *Soft Matter* **12**(7), 2086–2094 (2016)
- C. Zhou, S. Bette, U. Schnakenberg, Flexible and stretchable gold microstructures on extra soft poly(dimethylsiloxane) substrates. *Adv. Mater.* **27**(42), 6664–6669 (2015)
- Y. Zhu, F. Xu, Q. Qin, W.Y. Fung, W. Lu, Mechanical properties of vapor–liquid–solid synthesized silicon nanowires. *Nano Lett.* **9**(11), 3934–3939 (2009)



Hydrogen Embrittlement Cracking Produced by Indentation Test

8

Akio Yonezu and Xi Chen

Contents

Introduction	290
Evaluation of Threshold Stress Intensity Factor for HE Cracking	291
Materials and Experimental Methods	291
Experimental Results	292
Numerical Analysis	296
Mechanism of HE Cracking from Indentation Impression	302
Materials and Experimental Methods	302
Experimental Results	303
Discussion	305
Conclusion	311
References	312

Abstract

Indentation is a convenient method to evaluate mechanical properties of materials as well as to simulate contact fracture with locally plastic deformation. Indentation experiment has been widely used for brittle solids, including ceramics and glass, for evaluating the fracture properties. With the aid of computational framework, simulation of crack propagation (for quasi-static and dynamic impact) is conducted to characterize “brittleness” of materials. In this review, we explore the applicability of indentation method for hydrogen embrittlement

A. Yonezu (✉)

Department of Precision Mechanics, Chuo University, Tokyo, Japan

e-mail: yonezu@mech.chuo-u.ac.jp

X. Chen

Department of Earth and Environmental Engineering, Columbia Nanomechanics Research Center, Columbia University, New York, NY, USA

e-mail: xichen@columbia.edu

cracking (HEC). HEC is an important issue in the development of hydrogen-based energy systems. Especially high-strength steels tend to suffer from HE cracking, which leads to a significant decrease in the mechanical properties of the steels, including the critical stress for crack initiation and resistance to crack propagation. For such materials integrity for HEC, convenient material testing is necessary. In this review, the first part describes new indentation methodology to evaluate threshold stress intensity factor K_{ISCC} , and the latter one is investigation into HEC morphology due to residual stress produced by indentation impression. Our findings will be useful for predicting K_{ISCC} for HE instead of conventional long-term test with fracture mechanics testing. It will also indicate the stress criterion of HE cracking from an indentation impression crater, when the formed crater (for instance due to shot peening or foreign object contact) is exposed to a hydrogen environment.

Keywords

Indentation · Hydrogen embrittlement cracking · Fracture strength · Finite element method · Cohesive zone model · Residual stress · High-strength steel

Introduction

Hydrogen embrittlement (HE) cracking is an important issue in the development of hydrogen-based energy systems. High-strength steels tend to suffer from HE cracking, which leads to a significant decrease in the mechanical properties of the steels, including the critical stress for crack initiation and resistance to crack propagation. It is well known that Mode-I tensile stress is responsible for HE cracking (Gangloff 2003). In contrast, degradation of the mechanical properties upon compression loading (e.g., macroscopic hardness) never occurs in a hydrogen environment (Reddy et al. 1992). In order to evaluate the structural integrity of the material with respect to hydrogen embrittlement, tensile or bending loading is usually employed for mechanical testing, producing tensile stress for HE cracking. In contrast, we found that Vickers indentation caused HE cracking in high-strength steel (Yonezu et al. 2010). This is due to the fact that the tensile stress field that develops around the indentation impression reaches the critical value for crack nucleation. Since indentation testing is a convenient method to probe the mechanical properties of materials compared to other mechanical testing methods, such as fracture mechanics testing under tensile or bending loading, the ability to evaluate HE cracking upon indentation is of importance.

In addition, indentation test induces permanent impression crater, which is similar with shot peening. Peening techniques make the material harder but in addition also introduce compressive residual stress. It is well known that compressive stress is effective in preventing HE cracking. With multiple shot impacts on the surface, the surface hardened layer expands plastically and compressive residual stress develops (Kobayashi et al. 1998). However, the boundary of the peening area (outside the crater) sometimes shows tensile stress (Klemenza et al. 2009). This could lead to

mechanical degradation such as HE, stress corrosion cracking (SCC), and fatigue fracture. In fact, a single crater (permanent impression due to indentation) forms tensile stress around the impression, balanced by local compressive strain. Such a tensile stress may induce HE cracking, experienced in high-strength steels (Yonezu et al. 2010, 2012).

From the above engineering background, indentation method may bring the possibility to characterize HE cracking. In this review, the first topic is to explore the use of Vickers indentation testing for evaluating the susceptibility of steel to HE cracking. For this purpose, indentation test is applied to hydrogen-charged steel. The resistance to crack propagation, in particular, the threshold stress intensity factor (K_{ISCC}), is an important parameter for understanding the susceptibility to HE cracking. The second topic is to clarify HE cracking produced by residual stress of indentation impression. For this experiment, spherical impression crater is exposed in hydrogen environment, and HE cracking morphology and the mechanics are systematically investigated. This may become a potential fracture mode, when the steel forms a permanent crater on the surface due to surface treatment (e.g., shot peening) or other types of contact loading in a hydrogen environment. It is expected that the present knowledge is of importance and will significantly contribute to material/mechanical design in a hydrogen environment. In addition, the findings (of the first topic) allowed us to propose an alternative method to the fracture mechanics approach for HE cracking evaluation.

Evaluation of Threshold Stress Intensity Factor for HE Cracking

Materials and Experimental Methods

The material used in the present study was 18Ni maraging steel (350 ksi), which is a low-carbon martensitic steel. In order to achieve high strength, the steel was solution heat treated (820 °C for 3 h) followed by cooling in air to room temperature. Subsequently, the steel was aged at 500 °C for 4 h and air-cooled to room temperature. The mechanical properties of the steel are shown in Table 1, indicating that the yield stress and tensile strength were 2.40 and 2.45 GPa, respectively (Boyer and Gall 1985).

This steel has superior tensile strength, but is very susceptible to hydrogen embrittlement (Gangloff and Wei 1974; Pao and Wei 1977; Antlovich et al. 1980;

Table 1 Mechanical properties of 18Ni maraging steel (350ksi) employed in this study

Young's modulus E (GPa)	Poisson's ratio ν	Yield strength σ_y (GPa)	Tensile strength σ_B (GPa)	Work hardening coefficient n	Strength for HE cracking σ_{sc} (GPa)
210	0.3	2.4 ^a	2.45 ^a	0.025	0.65 ^b

^aBoyer and Gall (1985) Metals Handbook

^bYonezu et al. (2010)

Reddy et al. 1992; Gangloff 2003; Tsay et al. 2005, 2008). The specimens in the study were disk-shaped plates with a diameter of 20 mm and thickness of 1 mm. After mechanically and electrochemically polishing the specimen surface, cathodic charging was carried out in order to introduce hydrogen into the material in a buffer solution of sodium acetate (0.20 mol/L CH_3COOH + 0.17 mol/L CH_3COONa , pH = 4.7) with a current density of 3 A/m². In order to vary the hydrogen content in the steel, the duration of hydrogen charging was varied from 48 to 72 h.

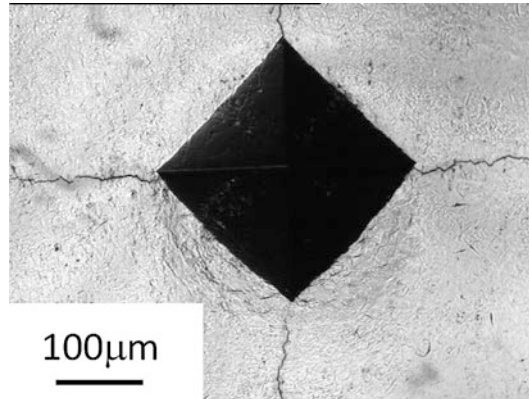
After cathodic charging, the total diffusible hydrogen content in the material was measured by thermal desorption spectroscopy (TDS-KU, ULVAC). Just after the cathodic charging, TDS measurements indicated C_{H} to be 33.7 ppm for 48-h charging and 49.1 ppm for 72-h charging. When the specimen with absorbed hydrogen is conditioned in air at room temperature, the hydrogen in the material diffuses out into the atmosphere as a function of the exposure time (degassing time). This suggests that C_{H} can be controlled by changing the degassing time. The discharged C_{H} as a function of conditioning time was also measured by TDS, indicating that the C_{H} for both specimens decreased with increasing degassing time (C_{H} was varied from 49.1 to almost 0 ppm). By referring to this hydrogen degassing behavior, the timing of the indentation test can be determined at a desired C_{H} in order to examine the relationship between the HE susceptibility and the C_{H} .

Indentation testing was carried out using a hydraulic servo-controlled fatigue testing machine equipped with a diamond Vickers indenter. The indenter impresses the specimen surface with up to a specific maximum indentation force at a rate of 1 N/s. Subsequently, the indentation force is reduced at a rate of 1 N/s until $F = 0$ N. The maximum force was varied from 100 to 300 N. The detailed conditions used will be described in each section below. After the indentation test, the impression morphology was observed using an optical microscope. These observations were conducted at least 1 h after the indentation test.

Experimental Results

Vickers indentation was applied to the specimen that had absorbed hydrogen as a result of cathodic charging for 72 h ($C_{\text{H}} = 49.1$ ppm). Figure 1 shows an optical micrograph of the impression and its surrounding showing that four cracks propagated from the corners of the impression. These are similar to surface cracks (radial and halfpenny-shaped cracks) observed in brittle materials subjected to Vickers indentation (Lawn et al. 1980; Niihara 1983; Cook and Pharr 1990). We previously reported that the cracks grew in a zigzag manner, indicating that they propagate along grain boundaries (Yonezu et al. 2010). Such intergranular cracks are often observed in maraging steel when subjected to hydrogen embrittlement (Gangloff 2003; Tsay et al. 2005, 2008). Furthermore, from observations of cross-sectional views, each crack was identified as a type of radial crack, which propagates along the radial direction from one corner of the indentation impression (Yonezu et al. 2010). In contrast, steel with no hydrogen absorption (as-received specimen: C_{H} less than 0.05 ppm) did not exhibit any cracking around the impression. We thus

Fig. 1 Micrographs of the indent impression produced in a hydrogen-charged specimen tested with $F_{\max} = 300$ N (hydrogen content is 49.1 ppm)



concluded that the crack formation in the present case was caused by hydrogen embrittlement (Yonezu et al. 2010).

Figure 2 shows the relationship between the surface crack length and the maximum indentation force. Here, the tested specimens were charged for 72 h, and all specimens had C_H of 49.1 ppm. The crack length (designed as c) is defined as the distance from the impression corner to the crack tip. The number of tests with $F_{\max} = 100, 200$ and 300 N were three, five and nine, respectively. Their standard deviations are also plotted in the figure. It should be noted that for one indentation impression, the number of cracks formed was usually four (such as seen in Fig. 1). However, in some impressions (especially those produced with a smaller $F_{\max} = 100$ N), the number of cracks tended to be two or three. Thus, the representative crack length for one test was assumed to be the averaged value, that is, the sum of the crack lengths divided by the number of cracks formed. Figure 2 reveals that the crack length c increases when F_{\max} is larger. Figure 2 also shows the results for half of the diagonal length in the impression (designed as a) as a function of F_{\max} , also indicating that the impression size is larger with increasing F_{\max} .

In order to investigate the effect of hydrogen content on the crack size, indentation tests were conducted at different times after cathodic charging for 72 h. The value of F_{\max} was set at 300 N for all tests. Figure 3 shows three micrographs of representative surface cracks in specimens with different hydrogen content which propagated from the impression corners; the tests were conducted at different discharge times (0 h (a), 43 h (b), and 119 h (c)) for hydrogen contents of 49.1, 26.9, and 19.5 ppm. Although F_{\max} was set at the same value (300 N) for all the tests, the crack length was strongly dependent on C_H , showing that the length became longer when C_H was larger.

Figure 4 shows the crack length as a function of C_H for all specimens. Here, the specimens with $C_H > 20$ ppm were obtained from 72-h cathodic charging, while the results for C_H around 5 ppm were obtained from the 48-h charging specimen. It can be seen that the crack length became larger with increasing C_H . Figure 4 also shows the Vickers hardness values measured from the impression size (marked

Fig. 2 Variations in crack length and impression size as a function of maximum indentation force

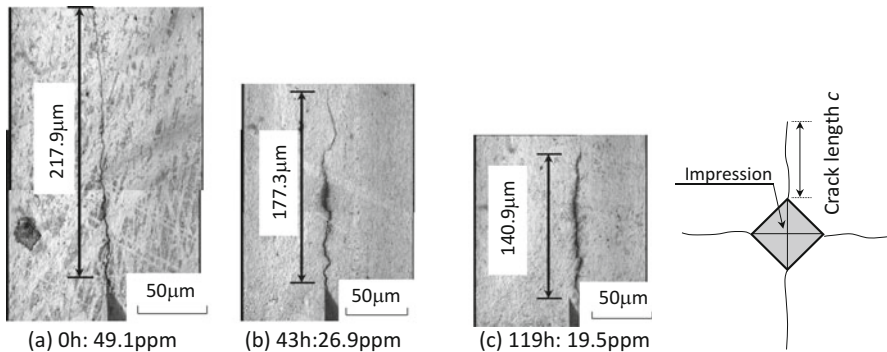
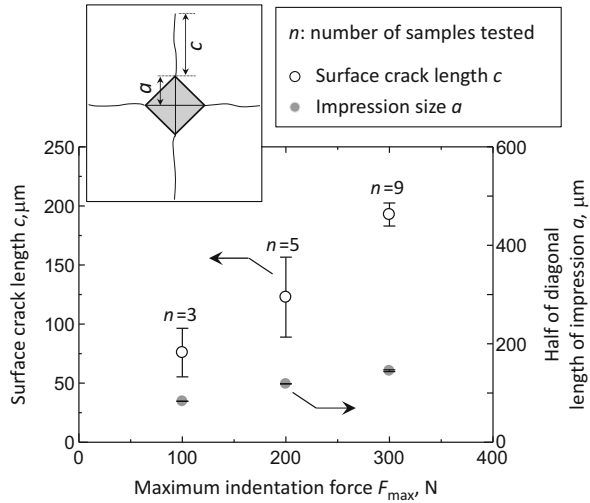


Fig. 3 Micrographs of the crack produced from an indent impression corner tested with $F_{max} = 300$ N. Three tests were conducted at different times after the hydrogen charging: (a) 0 h, (b) 43 h, and (c) 119 h. The hydrogen content was estimated to be (a) 49.1, (b) 26.9, and (c) 19.5 ppm

by a gray circle). The hardness did not depend on the C_H and showed an almost constant value equal to the hardness of the as-received steel with no hydrogen absorption ($HV = 675$) as indicated by the dotted line. Therefore, it is concluded that hydrogen in the steel influences the propagation of the Vickers indentation-induced crack, while the hardness related to the macroscopic elastoplastic properties in compression is not affected by the presence of hydrogen in the steel. Thus, hydrogen in the steel causes a decrease in the critical tensile strength, resulting in HE crack initiation and propagation. Propagation of the crack tip is expected to be suspended upon full unloading under the stress state, where the driving force of the crack propagation (due to the residual stress field from the indent formation) is equilibrated with the resistance to crack growth in the steel having hydrogen.

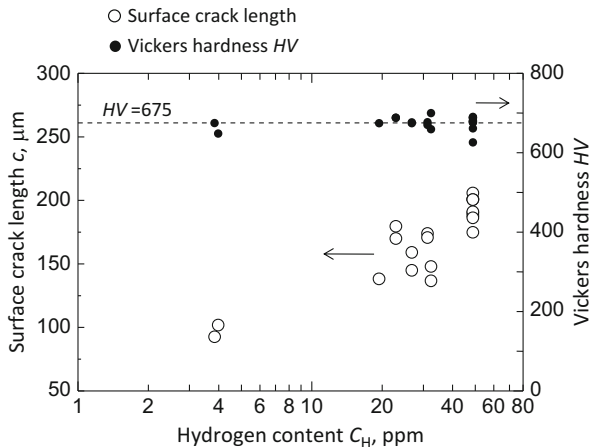


Fig. 4 Variations in radial surface crack length and Vickers hardness as a function of hydrogen content

In other words, the “final” crack length is dependent on resistance to HE crack propagation in the material.

Here, several questions regarding crack propagation arise, such as how the crack grows, what the growth rate is, and when crack propagation stops. In fact, K_{ISCC} tests based on fracture mechanics show that an HE crack starts to propagate when the stress intensity exceeds K_{ISCC} (the threshold resistance to HE crack growth), the growth rate rapidly increases, and the crack propagates stably at the constant crack growth rate (so-called plateau region). Finally, unstable fracture occurs at around K_{IC} . This implies that K_{ISCC} is much lower than K_{IC} . In other words, an HE crack cannot propagate when the stress intensity factor is lower than K_{ISCC} .

In our indentation tests, the total testing time (including both loading and unloading) was about 600 s, for the test condition of $F_{max} = 300$ N and loading rate of $dF/dt = 1$ N/s. If the crack initiates and propagates up to a length of 200 μm (see Fig. 3) during the indentation test, the average rate of crack growth may be about 3.33×10^{-7} m/s. In contrast, fracture mechanics tests yielded HE crack growth rates in the plateau region for 18Ni-300 (300 ksi grade) and 18Ni-250 (250 ksi grade) maraging steels of 5.0×10^{-5} (Sumitomo Precision Products CO. L.1992) and 1.11×10^{-7} m/s (Gangloff 2003), respectively. These values are of the same order as the average rate in the present study. This implies that the HE crack propagation due to indentation corresponds to the plateau region, indicating that the crack grows stably in the present tests.

We will next discuss when the crack stops propagating. Note that the measurement of crack length (in Fig. 3) in the present tests was conducted more than 1 h after the test. Considering the maximum crack length of about 200 μm for $F_{max} = 300$ N (see Fig. 3), assuming the crack continuously propagates until when the microscopic observations were conducted, the average crack growth rate is less

than 6.7×10^{-8} m/s. This value is smaller than the stable rate in the plateau region (5.0×10^{-5} and 1.11×10^{-7} m/s for 18Ni-300 and 18Ni-250 maraging steels as mentioned above), suggesting that the crack in the present study underwent stable propagation. In addition, when we continually observed the crack length up to 100 h after the test, the length did not change. In this case, the crack growth rate was less than about 1.1×10^{-12} m/s, since the measurement resolution of the optical microscope was less than $0.4 \mu\text{m}$. This rate is almost the same order as the threshold rate (2.8×10^{-12} m/s), when the value of K_{ISCC} was determined in the fracture mechanics test (Yamaguchi et al. 1997). Therefore, after full unloading, the driving force of the crack tip (stress intensity K) is considered to be less than the value of K_{ISCC} . Such a condition of $K \leq K_{\text{ISCC}}$ indicates that the HE crack stops propagating, resulting in the final crack length. In the next section, the stress field at the crack is computed to evaluate the threshold stress intensity of the final crack.

Numerical Analysis

Cohesive Zone Model

The stress field at the crack tip is a key issue, as expected from the above results, since the crack length (at full unloading) is dependent on the HE crack growth resistance of steel, K_{ISCC} . Note that the process of indentation loading/unloading remarkably changes the stress field, resulting in HE crack initiation and propagation (Yonezu et al. 2010). Therefore, such crack growth is required to be incorporated with stress analysis. We employed the finite element method (FEM) to compute the indentation stress field in conjunction with crack propagation.

One of the approaches to simulate crack propagation is the cohesive zone model (CZM). The CZM is theoretically well established and is proven to be applicable to both ductile and brittle materials (Barenblatt 1962; Tvergaard and Hutchinson 1992; Chandra et al. 2002; Xia et al. 2004; Hal et al. 2007; Olden et al. 2008). For instance, Mode-I fracture of a brittle coating on a silicon substrate (Xia et al. 2004), hydrogen embrittlement in a duplex stainless steel (Olden et al. 2008), and the interfacial fracture of an IC-interconnect (e.g., copper and low-K materials) (Hal et al. 2007) have been successfully investigated using the CZM.

Figure 5 shows a schematic of crack propagation incorporating the cohesive zone. The CZM essentially models the fracture process zone in a plane ahead of the crack tip. The zone is assumed to be subjected to cohesive traction. The model usually describes the gradual degradation of the adhesion between two regions along the crack propagation plane. The mechanical response of the cohesive zone obeys a traction-separation law that yields the relationship between the separation distance v of the two material faces at an interface and the traction stress σ acting between them. Figure 5 graphically explains the lumping of the nonlinear material response in the cohesive zone. By the action of the external force (i.e., remote stress/strain), the crack tip opens. The opening response at the actual crack tip (designated as v_{B}) obeyed the cohesive zone element with the traction-separation law, such as the

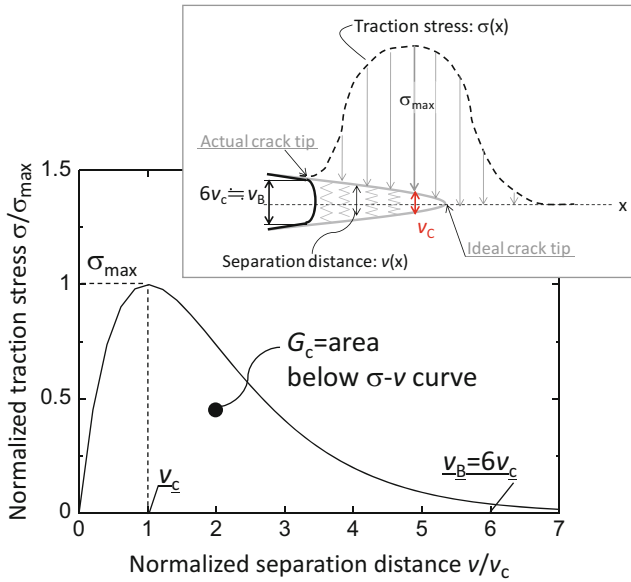


Fig. 5 Schematic of crack advancement with the cohesive zone model. The cohesive zone element having the traction-separation law exists ahead of the crack tip

bottom figure in Fig. 5. In the cohesive zone, after the acting stress σ exceeds the maximum cohesive stress σ_{max} (corresponding to the characteristic separation v_c), the resistance stress (traction force) dramatically decreases to almost zero with larger separation. This indicates that the cohesive zone ahead of the material crack tip opens, and crack propagation then occurs, resulting in the nucleation of new surfaces. Such a simulated crack is permitted to propagate along the specific direction along the cohesive zone element. As shown in Figs. 1 and 3, the surface crack propagates along the radial direction from the indent impression corner. The cohesive element is thus to be inserted in this plane.

Although numerous traction-separation laws for the cohesive zone element have been proposed (e.g., see reference (Chandra et al. 2002) for an overview), an exponential law, called the Smith-Ferrante type (Ortiz and Pandolfi 1999; Hal et al. 2007), is employed due to its simplicity. In fact, the exponential type has been used for various cases, including brittle Mode-I fracture, interfacial delamination, and hydrogen embrittlement cracking (Xia et al. 2004; Hal et al. 2007; Olden et al. 2008) as mentioned above. The constitutive behavior is expressed by the relationship between the normal traction stress σ and corresponding separation distance v across the cohesive element. This is given by

$$\sigma = \sigma_{max} \frac{v}{v_c} \exp\left(1 - \frac{v}{v_c}\right), \tag{1}$$

where σ_{\max} is the maximum cohesive stress and v_c is the characteristic opening displacement corresponding to σ_{\max} (see Fig. 5). Note that the tangential traction across this surface (i.e., shear component) is ignored in this study, since HE is caused by Mode-I loading and the driving force of the crack initiation and propagation is tensile stress (i.e., principal stress) (Yonezu et al. 2010). For the cohesive zone law, the normal work of fracture G_c corresponds to the area under the traction-separation law, as shown in this figure. It can be described as

$$G_c = \int_0^{\infty} \sigma(v) dv = \exp(1)\sigma_{\max}v_c. \quad (2)$$

Therefore, Eqs. (1) and (2) lead to

$$\sigma = \frac{G_c}{v_c^2} \exp\left(-\frac{v}{v_c}\right)v. \quad (3)$$

This indicates that the traction-separation curve depends on two variables, the critical energy release rate G_c and v_c (or σ_{\max}). When these two variables are known, the traction-separation curve can be drawn.

The critical energy release rate G_c is related to the critical stress intensity factor for crack growth of the material. When considering unstable crack growth, the fracture toughness K_{IC} can be obtained in the plane strain condition as shown below.

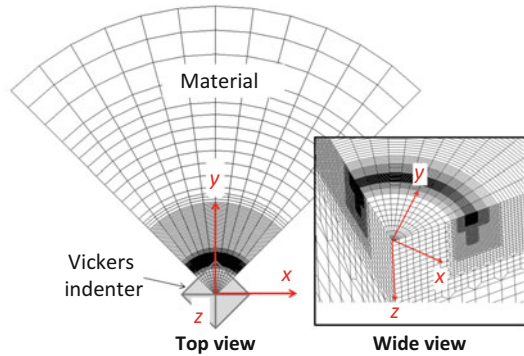
$$G_c = \frac{K_{IC}^2}{E} (1 - \nu^2) \quad (4)$$

Here, E and ν are the Young's modulus and Poisson's ratio, respectively. As mentioned above, the present study focuses on indentation HE cracking which is expected to be suspended under the condition of the driving force for crack propagation, $K \leq$ threshold stress intensity factor, K_{ISCC} . Thus, in Eq. (4), K_{IC} is hereafter replaced by K_{ISCC} .

Finite Element Method

A three-dimensional model of one-quarter of the specimen was created as shown in Fig. 6. The model contains more than 30,000 nodes and eight-node elements, and the part of the indenter contact and crack propagation (in the cohesive zone element) has a fine mesh. A mesh convergence study was carried out. The calculation was performed using commercially available FEM code (Marc and Mentat 2010.2). A rigid contact surface was used to simulate the rigid Vickers diamond indenter. Coulomb's law of friction was assumed with a friction coefficient of 0.15 (Bowden and Tabor 1950). The indenter penetrated to a maximum force of 300 N and was then withdrawn to zero. The mechanical properties of the as-received maraging steel, listed in Table 1, were used in the FEM computation because the macroscopic

Fig. 6 Stress computation of Vickers indentation with the three-dimensional FEM model



hardness, which is related to the macroscopic plastic properties in compression, does not change with hydrogen content (see Fig. 4).

In order to simulate crack propagation, a cohesive zone element was inserted in the y - z plane, where the crack extends from the corner of the impression. Thus, during indentation loading/unloading, the simulated crack propagates along the y - z plane, and such crack propagation obeys the traction-separation law. We next show computational results for propagation of various cracks. In this study, the actual crack tip is taken as the point where the normal displacement v equals $6v_c$, at which point the normal stress σ on the separated surface is essentially zero (Xia et al. 2004) (see Fig. 5). Since the tensile stress σ_{xx} is responsible for the HE crack propagation, we focused on the distribution of σ_{xx} in this study.

As an example case, the model with crack growth resistance G_c of 108.3 J/m^2 was computed. The maximum stress σ_{\max} in Eq. (1) was set to be 0.65 GPa , which was the critical stress for nucleating the present HE crack measured by Vickers indentation (Yonezu et al. 2010). Here, the value of G_c (108.3 J/m^2) for FEM can be converted to a K_{ISCC} value of $5.0 \text{ MPa m}^{1/2}$ from Eq. (4). Figure 7 shows a snapshot of the normal stress σ_{xx} distribution at maximum indentation force (a) and upon full unloading (b). For comparison, the model with *no* cohesive zone element was also computed, as shown in Fig. 7c, d. This model simulates the stress field due to indenter contact and does not induce any crack formation: the tensile stress σ_{xx} develops in the elastic field outside the finite plastic deformation region (local impression). Note that the magnitude of σ_{xx} in the fully unloaded condition (Fig. 7d) is higher than that at the maximum indentation force (Fig. 7c). This is because unloading reduces the elastic compressive stress due to the indenter contact, and therefore, σ_{xx} reaches a maximum value at full unloading.

In contrast, in Figs. 7a, b, the model with the cohesive zone element showed the release of tensile stress σ_{xx} (indicated by the dotted line in A) outside the impression (compressive plastic strain field). This is due to crack initiation and propagation. The FEM analysis showed the crack initiated at loading (Fig. 7a) and then propagated significantly under unloading (Fig. 7b). Finally, the crack length reaches a maximum upon full unloading.

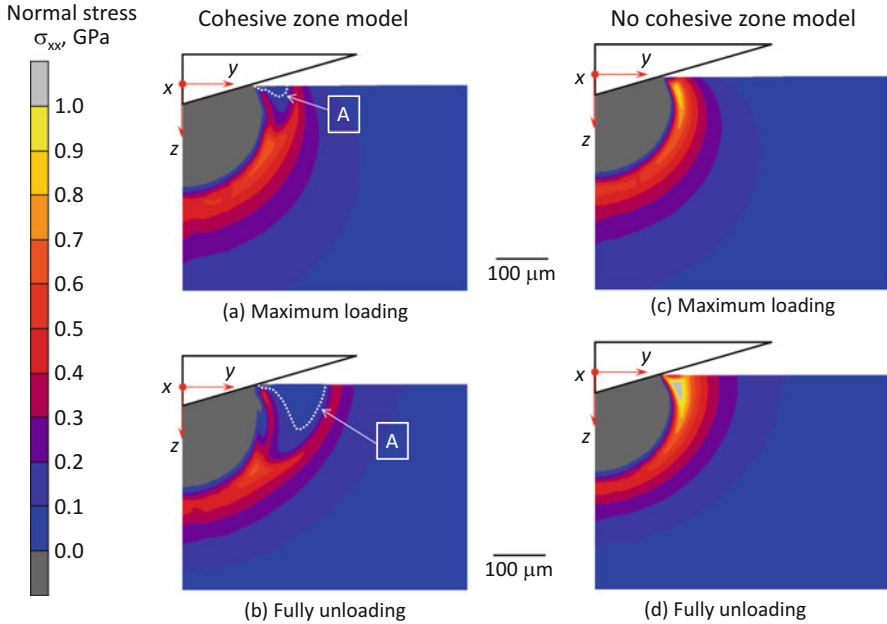


Fig. 7 Contour maps of σ_{xx} for the radial crack with the cohesive zone model (a) and (b) and without the cohesive zone element (c) and (d). Note Figs. (a) and (c) are shown at the maximum indentation force and (b) and (d) at the end of unloading

Crack Growth Resistance

To evaluate the crack growth resistance of the specimen with C_H of 49.1 ppm, an FEM study was carried out by changing K_{ISCC} from 3 to 8 MPa m^{1/2}, in order to analyze the relationship between K_{ISCC} and surface crack length. Here, the FEM model and its constitutive equation for the material were the same as the model described in Fig. 6, and the maximum indentation force was 300 N. Figure 8 shows the relationship between K_{ISCC} and the simulated surface crack length, c , indicating that the crack length is longer with a decrease in K_{ISCC} . The value of $K_{ISCC} \geq 8$ MPa m^{1/2} produced no crack. In this figure, the actual crack length obtained from the indentation experiment ($192.7 \pm 9.7 \mu\text{m}$ from Fig. 2) is also shown by the broken line. Here, the thin lines indicate the range obtained from the standard deviation of the experimental results. Compared with the experiments, K_{ISCC} can be estimated to be 4.41 MPa m^{1/2}, in the range from 4.28 to 4.55 MPa m^{1/2}. It should be emphasized that the variation in the estimated values is not very large, compared with those obtained by the fracture mechanics K_{ISCC} test (Floreen 1978).

To verify the estimated K_{ISCC} obtained from the test with $F_{\text{max}} = 300$ N, the other test results with indentation forces of 100 and 200 N were analyzed. Similar to Fig. 9, the experimental crack length (from Fig. 2) was used to estimate the values of K_{ISCC} . The estimated K_{ISCC} gives almost the same value (about 4 MPa m^{1/2} in Fig. 8). As shown in Fig. 3, the impression size as well as the crack length varied

Fig. 8 Variation in crack length during the indentation test computed by FEM incorporating the cohesive zone model

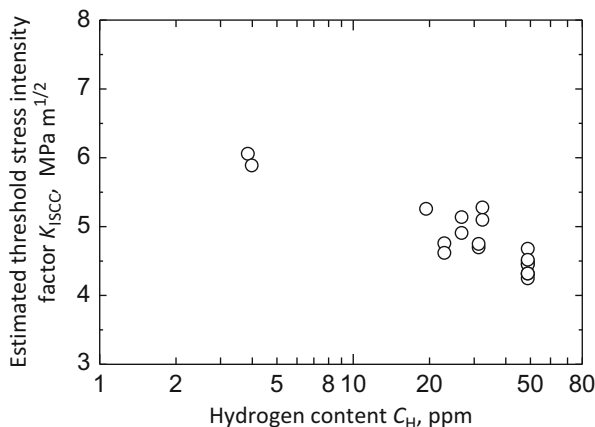
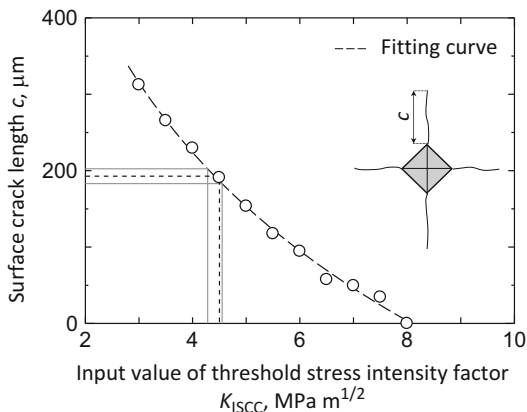


Fig. 9 Estimated threshold stress intensity factor K_{ISCC} as a function of the hydrogen content

depending on F_{max} . This indicates that the stress distribution (which is responsible for crack formation) is different. However, the estimated K_{ISCC} values were almost the same, suggesting that the present estimation is valid.

Figure 9 shows the K_{ISCC} value estimated by the present indentation method as a function of hydrogen content C_H . The K_{ISCC} values were calculated from the data as shown in Fig. 4. This suggests that the estimated values are a function of C_H , indicating that K_{ISCC} is lower with larger C_H . Such a trend was observed in previous studies on HE cracking evaluation for high-strength steels based on the fracture mechanics K_{ISCC} test (Yamaguchi et al. 1997; Gangloff 2003).

Table 2 summarizes the mechanical properties and threshold value of HE crack growth (i.e., threshold stress intensity factor for HE cracking, K_{ISCC}) obtained in the present study together with those for two different grades of maraging steels. The values of K_{ISCC} were obtained by the fracture mechanics test (K_{ISCC} test) in an H_2 environment for 18Ni-250 (250 ksi grade) and in 3.5% NaCl solution for 18Ni-300

Table 2 Comparisons of mechanical properties and threshold stress intensity factor K_{ISCC} among various 18Ni maraging steels

	18Ni maraging steel grade	Yield stress (GPa)	Tensile strength (GPa)	Fracture toughness ($\text{MPa m}^{1/2}$)	K_{ISCC} ($\text{MPa m}^{1/2}$)
This study	350ksi	2.4	2.45	40	4.24–6.05 ^c
Ref ^a	300ksi	2.07	2.17	75	5.5–11
Ref ^b	250ksi	1.65	1.72	120	20

^aSumitomo Precision Products CO, L. 1992, personal communication in 3.5% NaCl

^bGangloff (2003), in pressure H_2 ($P_{\text{H}_2} = 0.17 \text{ MPa}$)

^c K_{ISCC} estimated by the present study

(300 ksi grade) steel. The detailed conditions are described elsewhere (Sumitomo Precision Products CO, L. 1992, personal communication; Gangloff 2003). Since higher strength (higher yield/tensile strength) usually makes the HE susceptibility higher, the value of K_{ISCC} in 300 ksi grade steel is lower than for the 250 ksi grade steel. Our estimation order (4.41–6.05 $\text{MPa m}^{1/2}$) is almost the same as the K_{ISCC} values for the other steels (from 5.5 to 11 $\text{MPa m}^{1/2}$ for the 300 ksi grade, and 20 $\text{MPa m}^{1/2}$ for the 250 ksi grade), although the steel grade and hydrogen content are different. Furthermore, other maraging steels with a yield stress of about 2 GPa were reported to have K_{ISCC} of 10 $\text{MPa m}^{1/2}$ (McEvily 1990) and 7.6 $\text{MPa m}^{1/2}$ (Floreen 1978), indicating that the values are also close to our estimated values. Thus, this suggests that our method based on indentation can indeed be used to evaluate K_{ISCC} .

Compared with the indentation tests, the fracture mechanics K_{ISCC} test requires much more effort for sample preparation and setting up of the testing method, etc. For instance, the value of K_{ISCC} is usually obtained when crack growth has stopped, with decreasing driving force for crack growth. To recognize the termination of crack propagation, the experiment needs to last a minimum of 1000 h. Thus, the K_{ISCC} test usually requires a long time. In contrast, our method based on indentation saves time and potentially is an alternative technique.

Mechanism of HE Cracking from Indentation Impression

Materials and Experimental Methods

The above section is to investigate whether indentation cracking occurs or not, when the indentation loading is applied to hydrogen-charged steel. On the contrary, indentation cracking from permanent impression due to hydrogen absorption (i.e., when the impression crater is exposed in hydrogen environment) is clarified in this section. We used the same steel, i.e., 18Ni maraging steel (called 350 ksi (Boyer and Gall 1985)), which is a low-carbon martensitic steel. This section used a spherical diamond indenter with a 400 μm radius, which is quite a bit larger than the grain

size of the maraging steel (about 8 μm observed in Yonezu et al. 2010) used in this work. The typical shot peening crater has a diameter of several hundred μm , and in this case, we can ignore the effect of individual crystal plasticity on the stress analysis (based on continuum mechanics). Thus, a relatively large impression crater is required. The maximum indentation forces are 100, 200, and 300 N. The loading rate is set to 1 N/s for all tests, thereby simulating quasi-static contact loading.

After the indentation test to create a permanent impression, cathodic charging was carried out to introduce hydrogen into the material using a phosphate buffer solution (2.6 w/v% K_2HPO_4 , 0.2 w/v% NaOH, pH = 6.5) with a current density of 5 A/m^2 . In order to vary the hydrogen content in the steel, the duration of hydrogen charging was varied, and two durations of 6 and 48 h were used. The impression morphology was observed using an optical microscope (BXM-N33 M, Olympus Corp.). After the hydrogen charging, the sample was cleaned in ethanol in an ultrasonic cleaner (UT-206, SHARP Corp) for more than 1 h. Some specimens were mechanically cleaned along with fine polishing of buffing in order to remove the oxide film (formed during hydrogen charging) for clear observations. We were able to confirm that the length of the crack did not change before and after the fine polishing to remove the oxide film. Observations of the impression were conducted at least 1 h after the hydrogen charging. We confirmed that the hydrogen crack completely stops propagating after hydrogen charging.

Experimental Results

Indentation tests with different maximum indentation forces of $F_{\text{max}} = 100, 200$ and 300 N were carried out. Cathodic charging of the samples was then carried out for 6 and 48 h. For each test, about 30 impressions were performed. Figure 10 shows optical micrographs of representative indentation impressions and the surrounding area for hydrogen charging of 48 h. These figures separately show the results as a function of hydrogen charging time and F_{max} value. In Fig. 10a of $F_{\text{max}} = 100$ N test, the impressions show short cracks (less than about 50 μm) from the rim. Indeed, it is difficult to observe these, since the crack width is very narrow. In Fig. 10b of $F_{\text{max}} = 200$ N tests, however, long cracks are clearly observed from the rim. The cracks propagate radially from the impression, indicating a radial crack. Short cracks (similar to Fig. 10a) are also observed in these samples. Figure 10c (of $F_{\text{max}} = 300$ N test) show the same trend, indicating two types of cracks. It should be noted that the steel with no hydrogen absorption (before hydrogen charging) did not exhibit any cracking around the indentation impression (Yonezu et al. 2010) (as explained before). The present results indicate that permanent indentation impressions in the steel used absorbed hydrogen-induced short cracks and sometimes long cracks. Although the pictures of 6-h charging tests are omitted, it is found that the crack morphology also changes depending on the hydrogen content and applied indentation force (Niwa et al. 2015).

Figure 11 shows the frequencies of different crack lengths for each test condition. Similar to the series, Figs. 10 and 12 show results separately for different maximum

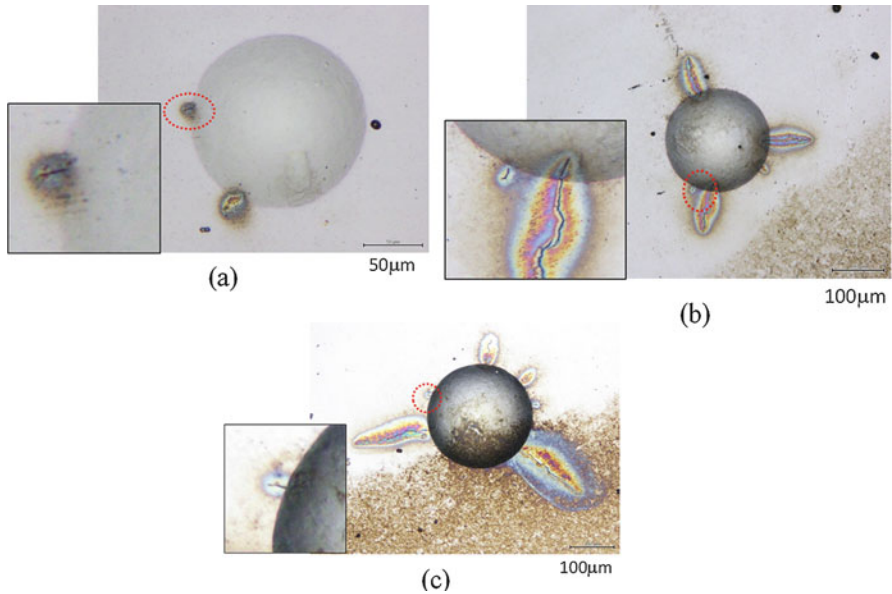


Fig. 10 Micrographs of an impression crater and its surrounding area. The hydrogen charging time is 48 h. The maximum indentation force $F_{max} = 100$ N for (a), 200 N for (b), and 300 N for (c)

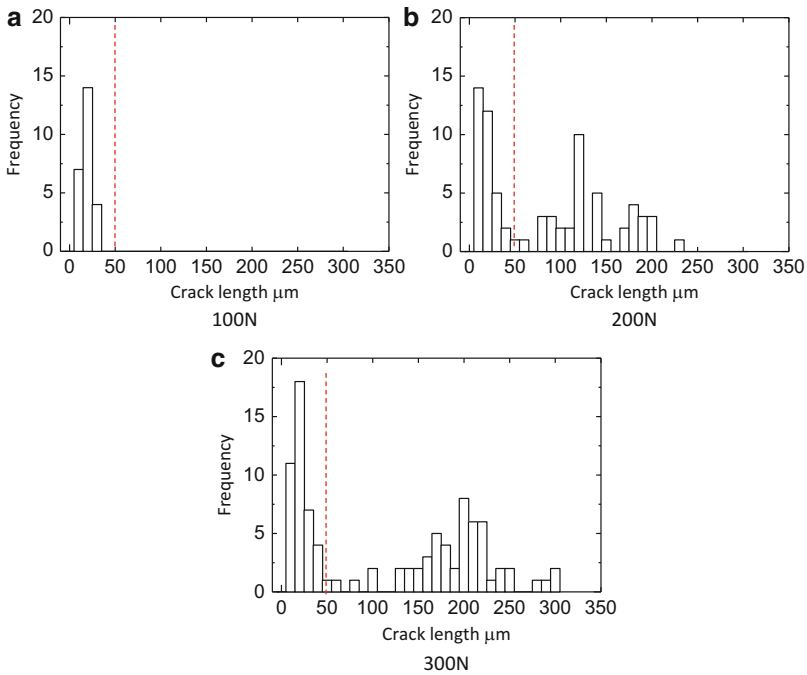


Fig. 11 Frequency of crack lengths from the indentation impressions for $F_{max} = 100$ N (a), 200 N (b), and 300 N (c) (H_2 charging time, 48 h)

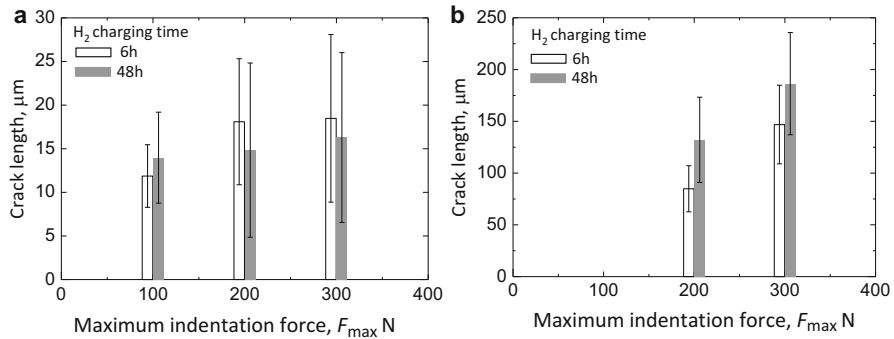


Fig. 12 Changes in crack length (for short cracks (a) and long cracks (b)) as a function of maximum indentation force

indentation force F_{max} (100, 200, and 300 N). For the $F_{\text{max}} = 100$ N test (Fig. 11a), short cracks with length less than 50 μm were observed for both charging times. On the other hand, the tests for $F_{\text{max}} = 200$ and 300 N showed two peaks in the distributions, where both short and long cracks exist. As also seen in Fig. 10, the two types of cracks (short and long cracks) can be distinguished, and the boundary between them is set to 50 μm , as shown by the dashed vertical line in Fig. 11.

Subsequently, we measured the crack length for both crack types, as shown in Fig. 12. The short crack length (Fig. 12a) seems to be independent of F_{max} and hydrogen charging time; however, the data shows large scatter. In contrast, Fig. 12b shows that the long crack length has strong dependency on F_{max} . In addition, the crack length is dependent on the hydrogen charging time, and the crack length is longer with larger charging time. The stress distributions and criteria were clarified in the study using finite element simulations, as described in the next section.

Discussion

Mechanism of Crack Propagation

In order to investigate crack formations of Fig. 10, the cohesive zone model (CZM) was employed in FEM to compute the stress field in conjunction with radial crack propagation. As mentioned above (section “Cohesive Zone Model”), this study employed an exponential law (called the Smith-Ferrante law) due to its simplicity. This law requires two independent material parameters, i.e., the maximum stress σ_{max} and the crack growth resistance K_C . The σ_{max} roughly corresponds to the critical stress for crack nucleation. As discussed before, the critical stress (σ_{max}) is estimated to be 0.65 GPa in the present steels which suffer from hydrogen embrittlement (Yonezu et al. 2010, 2012, 2015). In contrast, it is well known that K_C (which is similar to K_{ISCC} as discussed earlier) is strongly dependent hydrogen accumulation (Yamaguchi et al. 1997; Yonezu et al. 2012), and thus K_{ISCC} is not known. In fact, this study explores the actual K_{ISCC} with comparison between the

FEM computation and experiment (see later, section “[Crack Length and Stress Intensity Factor](#)”).

We created a three-dimensional FEM model of one-quarter of the specimen (similar with Fig. 6). One CZM element for crack propagation is inserted along the center line (at a 45-degree angle from the side surface). It is noted that the situation of crack propagation is different from Fig. 6, because HE crack propagates from permanent impression crater (i.e., crack propagation starts “post” indentation). A permanent impression is first made by spherical indentation with $F_{\max} = 100, 200$ and 300 N. After full unloading, the cohesive zone element operates, such that the crack starts propagating. We simulated crack propagation by the residual stress developed around the impression. To make the indenter impression, a rigid spherical indenter was used in a similar manner to that shown in Fig. 6. The mechanical properties used for the FEM are listed in Table 1.

Figure 13 shows representative contour maps of the crack nucleation area upon full unloading. The figure shows cross-sectional views along the crack propagation. For all models, the input value of the stress intensity factor was set to $3.5 \text{ MPa m}^{1/2}$, as a representative case. The maximum indentation force was changed to $F_{\max} = 300$ N, 200 N, and 100 N, and the results are shown separately in Fig. 13a, b, c, respectively. In the figures, the yellow area indicates the crack nucleation area (i.e., the crack surface). Here, the yellow region defines the crack nucleation area where the CZM interface completely opens, and there is no traction force between the CZM interfaces. It is found that the crack length strongly depends on F_{\max} . In particular, the test with $F_{\max} = 100$ N produces very small cracks, whose length is about $29 \text{ }\mu\text{m}$ (categorized as a “small crack” in this study). On the other hand, the larger F_{\max} tests ($F_{\max} = 200$ and 300 N) produced long cracks, which were very deep. This computational result corresponds to the experimental results in Figs. 10 and 11.

However, the larger F_{\max} tests ($F_{\max} = 200$ and 300 N) not only formed long cracks but also small cracks (even if the circumferential stress surrounding the indenter impression was uniform). The question arises as to why two different crack types exist (i.e., the larger F_{\max} produces both short and long cracks). The mechanism of such a “multiple crack” scenario was investigated as discussed in the next section.

Mechanism of Multiple Crack Formation

Multiple crack formation with two types of cracks with different lengths was observed, as discussed above. We created a simple FEM model for simulating such a multiple crack propagation. Figure 14 shows a three-dimensional model with half size. Two CZM elements are inserted along the radial direction, so that it can simulate radial crack propagation. Similar to Fig. 13, when the indenter is completely withdrawn, the CZM element starts operating, such that we can simulate crack propagation from the residual stress around the indenter impression. To make the indenter impression, a rigid spherical indenter was used, and the maximum force F_{\max} is 300 N. The three-dimensional half model in Fig. 14 comprises 53,200 eight-node elements. The mechanical properties used for the FEM are listed in Table 1.

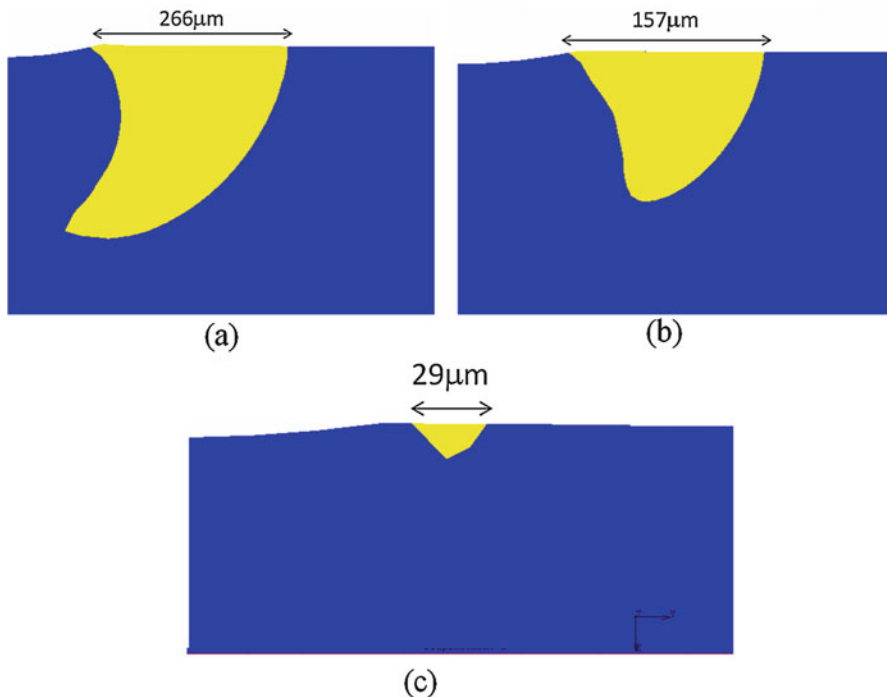


Fig. 13 Contour map of crack nucleation area in the model with $K = 3.5 \text{ MPa m}^{1/2}$. $F_{\text{max}} = 300 \text{ N}$ (a), 200 N (b), and 100 N (c)

As shown in Fig. 14, two CZM elements (designated as CZM① and CZM②) are inserted at a certain angle θ . We investigated the relationship between the crack propagation behavior and the crack angle θ , which can be influenced by the interaction of multiple cracks. Indeed, when a crack propagates, the stress around the crack is released, and it may be difficult for a new crack to propagate around that area. However, with increasing distance from the propagated crack, the stress recovers to the original state, resulting in new crack propagation. It should be noted that the distance in this case is the angle in the circumferential direction, and in the present study, this crack angle was set from 30 to 90 degrees, as shown in Fig. 14.

As mentioned above, we experimentally observed the occurrence of both long and short cracks together. In the present case, it seemed that a long crack was generated first, resulting in the surrounding stress field becoming weak. Then, the second crack near the long crack becomes short, while the crack far from the first long one becomes long (i.e., it recovers to the length of the first long crack). The crack length might be dependent on the angle from the first long crack. Based on this assumption, after full unloading, CZM① operates first, followed by the operation of CZM②. Such a “delay” in the computation process is controlled by

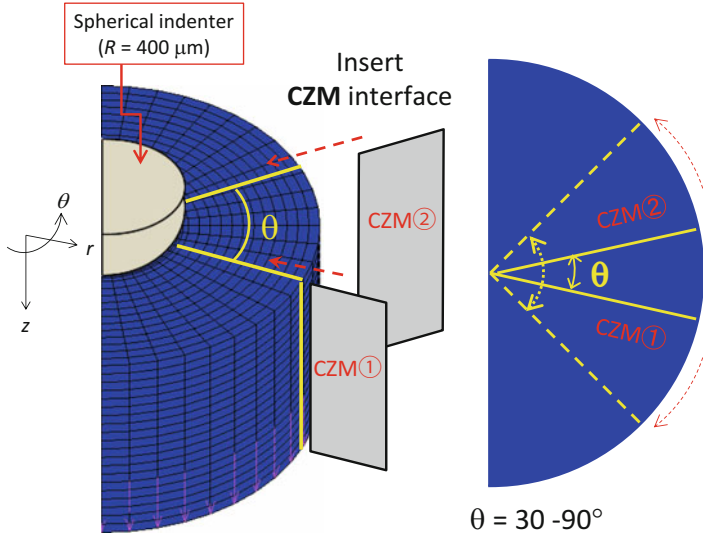


Fig. 14 Three-dimensional FEM model (half model) with cohesive zone element (CZM) for simulating radial crack propagation. This model simulates multiple crack propagation using CZM and CZM whose angle is changed from 30 to 90 degrees

the step increment in the FEM analysis. Of course, if CZM① and CZM② are generated at the same time, their lengths must be the same. In reality, however, it is unlikely to guess that all cracks are generated at the same time during the hydrogen charging, since crack initiation is strongly dependent on both inherent defects (which are present inhomogeneously) and the hydrogen accumulation state. Thus, we hypothesize that multiple cracks initiate separately at different times, i.e., CZM① is the “primary crack” and CZM② is the “secondary crack” (in situ monitoring of hydrogen cracking cannot yet be conducted in this study). Note that the material parameters of CZM① and CZM② are identical.

Using the model shown in Fig. 14, we investigated changes in the crack length of CZM① and CZM② as a function of the angle between the two cracks. As a representative case, the threshold stress intensity factor of the crack growth K_{ISCC} was set to $4.5 \text{ MPa m}^{1/2}$ for the CZM parameter. Figure 15 shows the results for the computed crack length for two CZM elements. The figure shows top views of the contour map, with the crack propagation area shown as yellow. Figure 15a shows the result for the 30-degree angle, while the result for the 90-degree angle is shown in Fig. 15b. When the angle is 90 degrees (larger angle), the crack lengths of CZM① and CZM② are identical. In contrast, when the angle is 30 degrees, the secondary crack (CZM②) becomes significantly shorter.

We next investigated how the crack length (especially of CZM②) is changed by the angle between the cracks. The results are shown in Fig. 16. It can be seen that the primary crack length of CZM① has little dependence on the angle, while

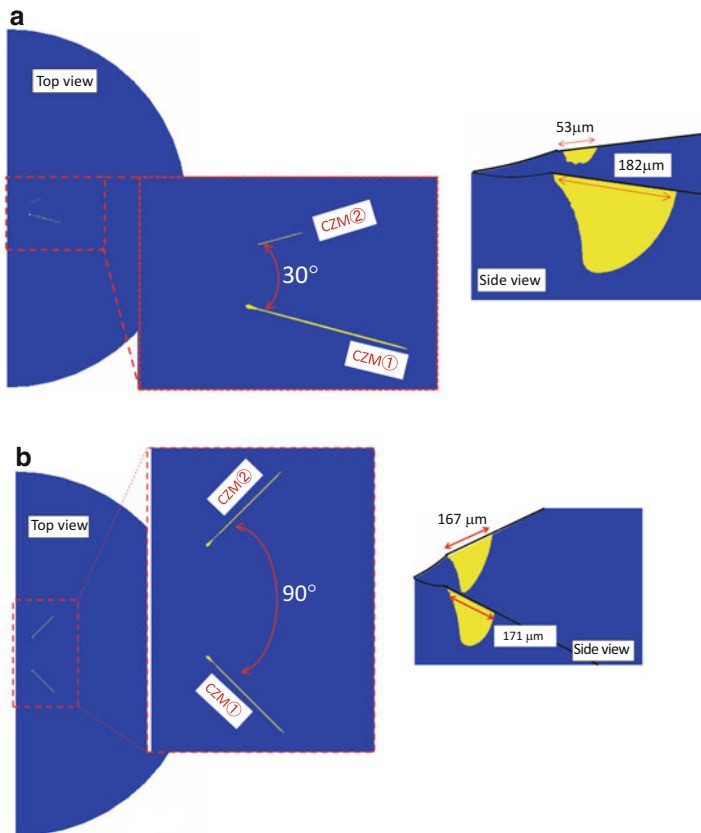


Fig. 15 Contour map of crack nucleation area, showing the multiple crack propagation (crack angle of 30 degrees (a) and 90 degrees (b))

the secondary crack (CZM②) is strongly dependent on the angle. At an angle of 90 degrees, the length of CZM② agrees well with that of CZM①. This suggests that at 90 degrees, multiple cracks do not interact with each other. However, at a smaller angle (when the position of CZM② is closer to CZM①), the length of CZM becomes significantly shorter. Finally, the length of CZM at about 30 degrees becomes less than 50 μm, and angles smaller than 30 degrees do not produce any cracks in this model. This might be due to the fact that the driving force for the CZM crack becomes small owing to propagation of the primary crack (CZM). Thus, the effect of primary crack generation becomes significant when the crack angle is small. It can therefore be concluded that the above discussion is the mechanism for the appearance of a short crack together with a long crack. It should be noted that this trend (crack length vs. crack angle) is dependent on the crack growth resistance, K_{ISCC} . It is expected that the case of smaller K_{ISCC} leads to a smaller crack angle, while a larger K_{ISCC} leads to a larger angle.

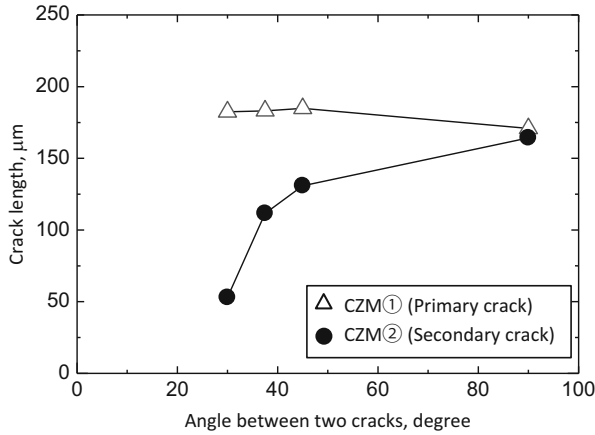


Fig. 16 Simulated crack length with respect to the angle between the two cracks (CZM and CZM)

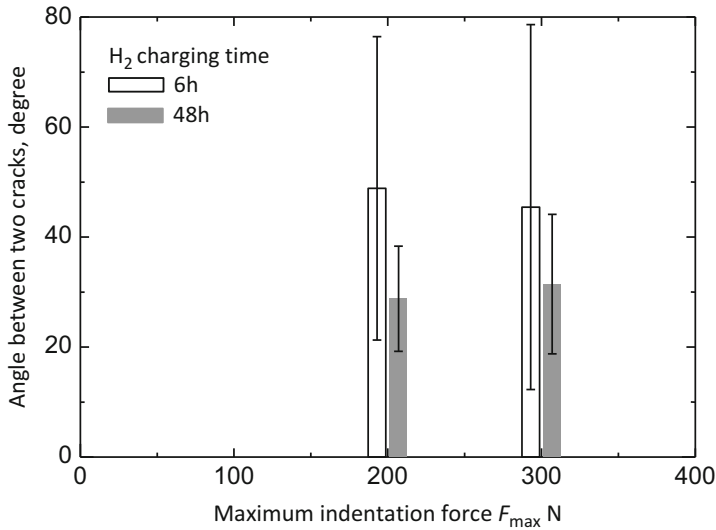


Fig. 17 Experimental result of the angle between the two cracks (short and long cracks)

Experimentally, we measured the angle between the long and short cracks for all indenter impressions. In this study, we only measured the minimum angle between the long and short cracks as the crack angle θ . The results are shown in Fig. 17 as a function of the applied F_{max} and the hydrogen charging time. Although the data show large scatter, the average crack angle θ was about 45 degrees for hydrogen charging for 6 h and 30 degrees for 48-h hydrogen charging. Although θ is dependent on K_{ISCC} , the experimental data are similar to the computational results in Fig. 17, especially for the 48-h charging time (the angle is about 30 degrees).

Crack Length and Stress Intensity Factor

We finally discuss the actual crack length from experimental HE cracking based on stress analysis, especially the stress intensity factor around the crack tip. Here we focus on the “long crack” (see Fig. 16), since it is not influenced by other crack generation. Similar with Fig. 8 of Vickers indentation, we investigated the effect of the input value of the stress intensity factor (crack growth resistance) K_{ISCC} on the crack length. We conducted several computations in this study with various values of the crack growth resistance (stress intensity factor K was changed from 3.0 to 5.5 MPa m^{1/2} in steps of 0.5 MPa m^{1/2}) for both the $F_{max} = 200$ N and 300 N tests. As expected, the crack length is strongly dependent on the input K_{ISCC} , indicating that the crack length is longer when K_{ISCC} is smaller (like Fig. 8). In addition, a strong dependence of F_{max} on the crack length was observed, i.e., with larger F_{max} , the radial crack becomes longer. This relationship can yield the actual K_{ISCC} when the crack length is known. Therefore, the experimental crack length is substituted in this relationship in order to estimate the actual K_{ISCC} .

The estimated values are investigated as a function of the maximum indentation force, F_{max} . This indicates that the estimated K_{ISCC} depends on the hydrogen charging time (i.e., hydrogen content), while it is not dependent on F_{max} . For the hydrogen charging time of 6 h, K_{ISCC} is estimated to be 4.9 MPa m^{1/2}, while for the charging time of 48 h, K_{ISCC} is 4.3 MPa m^{1/2}. As mentioned in section “Crack Growth Resistance,” the K_{ISCC} via Vickers indentation is estimated from 4 to 6 MPa m^{1/2} (depending on the hydrogen content). As shown in Table 2, furthermore, fracture mechanics tests for HE cracking reported the values of K_{ISCC} to be in the range of 5.5 to 11 MPa m^{1/2} (although the steel grade and hydrogen content in these tests are different from those in the present study). These values are in reasonable agreement with our estimation, suggesting that the present crack length (long crack) is governed by the threshold stress intensity factor K_{ISCC} for HE. Thus, the present findings may be useful for prediction of HE crack morphology from indentation impression.

Conclusion

This study investigated indentation-induced cracking in a maraging steel due to hydrogen embrittlement (HE). This review addressed two topics. The first one is that Vickers indentation was applied to the hydrogen-charged steel, such that we explored new evaluation method for HE susceptibility. The second topic is to investigate HE crack morphology from the residual stress that develops around a spherical indentation impression crater.

In the first topic, it is discovered that Vickers indentation induces hydrogen embrittlement cracking. The surface crack length is dependent on the hydrogen content in the steel. Since the stress intensity factor is equilibrated with the threshold stress intensity factor, K_{ISCC} , the stress field upon indentation in conjunction with crack growth was computed by the finite element method (FEM) incorporating a cohesive zone model. It was found that the simulated crack starts to grow under

loading and then propagates under unloading. The crack length finally reaches a maximum value at full unloading. The FEM with different values of the crack growth resistance (K_{ISCC}) was used to obtain the relationship between K_{ISCC} and surface crack length. By substituting the actual crack length in the relationship, K_{ISCC} could be estimated. The estimated values were found to agree with the K_{ISCC} values measured by fracture mechanics testing (K_{ISCC} test). Furthermore, our method showed a dependence of hydrogen content on K_{ISCC} . This trend is observed when fracture mechanics K_{ISCC} tests are applied to high-strength steel in a hydrogen environment. Such a HE evaluation is usually conducted by K_{ISCC} tests that require significant effort for specimen preparation and setting up of testing, etc. However, the indentation-based evaluation method proposed in this study is very simple and convenient. The present approach is an alternative for evaluation of HE cracking.

The second topic clarified the stress criterion of hydrogen embrittlement cracking (i.e., HE crack initiation and propagation) from the residual stress of indentation impression. Namely, when spherical impression crater is exposed in hydrogen environment, how HE cracking propagates from the crater was investigated. The HE crack propagates radially, and the crack length is dependent on maximum indentation force and critical crack growth resistance K_{ISCC} (which is changed by the hydrogen content). In addition, a larger indentation force results in a complicated HE fracture, showing both long and short cracks. In this case, the appearance of the short cracks is affected by stress release of a neighboring long crack. These findings in our systematic investigation may be useful to predict damage of HE cracks, produced by residual stress in the permanent impression. Especially, this will be useful to determine how the size of an impinging foreign object (indenter size and shape) and the loading magnitude (maximum indentation force) induce HE damage, including crack length. This may become an indicator for HE upon permanent impression crater/dimple formation, such as by shot peening and localized contact loading.

Acknowledgment We would like to thank Professor Kohji Minoshima (Osaka University) for his guidance. This work is supported by JSPS KAKENHI (Grant No. 22760077 and 26420025) from the Japan Society for the Promotion of Science (JSPS) and Research Grant of Suga Weathering Technology Foundation (SWTF).

References

- S.D. Antlovich, T.R. Risbeck, et al., The effect of microstructure on the fracture toughness of 300 and 350 grade maraging steels. *Eng. Frac. Mech.* **13**, 717–739 (1980)
- G.I. Barenblatt, Mathematical theory of equilibrium cracks. *Adv. Appl. Mech.* **7**, 56–129 (1962)
- F.P. Bowden, D. Tabor, *The Friction and Lubrications of Solids* (Oxford University Press, Oxford, 1950)
- H.E. Boyer, T.L. Gall, *Metals Handbook Desk Edition Metals Park* (American Society for Metals, Ohio, 1985)
- N. Chandra, H. Li, et al., Some issues in the application of cohesive zone models for metallic-ceramic interfaces. *Int. J. Solids Struct.* **39**, 2827–2855 (2002)

- R.F. Cook, G.M. Pharr, Direct observation and analysis of indentation cracking in glasses and ceramics. *J. Am. Ceram. Soc.* **73**, 787–817 (1990)
- S. Floreen, Maraging steels/metals handbook 9th edn. vol 1, in *Properties and Selection: Irons and Steels Metals Park*, (American Society for Metals, Ohio, 1978)
- R.P. Gangloff, *Hydrogen Assisted Cracking of High Strength Alloys* (Elsevier Science, New York, 2003)
- R.P. Gangloff, R.P. Wei, Gaseous hydrogen assisted crack growth in 18 nickel maraging steels. *Scri. Metall.* **8**, 661–668 (1974)
- B.A.E. Hal, R.H.J. Peerlings, et al., Cohesive zone modeling for structural integrity analysis of IC interconnects. *Microelectron. Reliab.* **47**, 1251–1261 (2007)
- M. Klemenza, V. Schulzea, et al., Application of the FEM for the prediction of the surface layer characteristics after shot peening. *J. Mater. Process. Technol.* **209**, 4093–4102 (2009)
- M. Kobayashi, T. Matsui, et al., Mechanism of creation of compressive residual stress by shot peening. *Int. J. Fatigue* **20**(5), 351–357 (1998)
- B.R. Lawn, A.G. Evans, et al., Elastic/plastic indentation damage in ceramics: the median/radial crack system. *J. Am. Ceram. Soc.* **63**, 574–580 (1980)
- A.J. McEvily, *Atlas of Stress Corrosion and Corrosion Fatigue Curves* (ASM International, Ohio, 1990)
- K. Niihara, A fracture mechanics analysis of indentation-induced Palmqvist cracks in ceramics. *J. Mater. Sci. Lett.* **2**, 221–223 (1983)
- M. Niwa, T. Shikama, et al., Mechanism of hydrogen embrittlement cracking produced by residual stress from indentation impression. *Mater. Sci. Eng. A* **624**, 52–61 (2015)
- V. Olden, C. Thaulow, et al., Application of hydrogen influenced cohesive laws in the prediction of hydrogen induced stress cracking in 25% Cr duplex stainless steel. *Eng. Fract. Mech.* **75**, 2333–2351 (2008)
- M. Ortiz, A. Pandolfi, Finite-deformation irreversible cohesive elements for three-dimensional crack-propagation analysis. *Int. J. Num. Methods Eng.* **44**, 1267–1282 (1999)
- P.S. Pao, R.P. Wei, Hydrogen assisted crack growth in 18Ni (300) maraging steel. *Scri. Metall.* **11**(6), 515–520 (1977)
- K.G. Reddy, S. Arumugam, et al., Hydrogen embrittlement of maraging steel. *J. Mater. Sci.* **27**, 5159–5162 (1992)
- L.W. Tsay, Y.F. Hu, et al., Embrittlement of T-200 maraging steel in a hydrogen sulfide solution. *Corros. Sci.* **47**, 965–976 (2005)
- L.W. Tsay, H.L. Lu, et al., The effect of grain size and aging on hydrogen embrittlement of a maraging steel. *Corros. Sci.* **50**, 2506–2511 (2008)
- V. Tvergaard, J.W. Hutchinson, The relation between crack-growth resistance and fracture process parameters in elastic plastic solids. *J. Mech. Phys. Solids* **40**, 1377–1397 (1992)
- Z. Xia, W.A. Curtin, et al., A new method to evaluate the fracture toughness of thin films. *Acta Mater.* **52**, 3507–3517 (2004)
- Y. Yamaguchi, H. Nonaka, et al., Effect of hydrogen content on threshold stress intensity factor in carbon steel in hydrogen-assisted cracking environments. *Corrosion-NACE* **53**(2), 147–155 (1997)
- A. Yonezu, M. Arino, et al., On hydrogen-induced Vickers indentation cracking in high-strength steel. *Mech. Res. Comm.* **37**, 230–234 (2010)
- A. Yonezu, T. Hara, et al., Evaluation of threshold stress intensity factor of hydrogen embrittlement cracking by indentation testing. *Mater. Sci. Eng. A* **531**(1), 147–154 (2012)
- A. Yonezu, M. Niwa, et al., Characterization of hydrogen-induced contact fracture in high strength steel. *J. Eng. Mater. Technol.* **137**, 021007 (2015)



Continuous Stiffness Measurement Nanoindentation Experiments on Polymeric Glasses: Strain Rate Alteration

9

George Z. Voyiadjis, Leila Malekmoitei, and Aref Samadi-Dooki

Contents

Introduction	316
Materials and Methods	318
Sample Preparation	318
Nanoindentation Analysis	318
The Indentation Strain Rate	320
Experimental Procedure	321
Results and Discussion	321
Variation of the \dot{P}/P in the Course of an Indentation	321
Variation of the Strain Rate in the Course of an Indentation	322
Indentation Size Effect	326
Conclusions	331
References	331

Abstract

In many studies using continuous stiffness measurement (CSM) nanoindentation technique, it is assumed that the strain rate remains constant during the whole experiment since the loading rate divided by the load (\dot{P}/P) is considered as a constant input parameter. Using the CSM method, the soundness of this assumption in nanoindentation of polymeric glasses is investigated by conducting a series of experiments on annealed poly(methyl methacrylate) (PMMA) and polycarbonate (PC) at different set \dot{P}/P values. Evaluating the variation of the

G. Z. Voyiadjis (✉) · L. Malekmoitei
Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge,
LA, USA
e-mail: voyiadjis@eng.lsu.edu; lmalek1@lsu.edu

A. Samadi-Dooki
Computational Solid Mechanics Laboratory, Department of Civil and Environmental
Engineering, Louisiana State University, Baton Rouge, LA, USA
e-mail: asamad3@lsu.edu

actual \dot{P}/P value during the course of a single test shows that this parameter varies intensely at shallow indentation depths, and it reaches a stabilized value after a significant depth which is not material dependent. In addition, the strain rate variation is examined through two methods: first, using the definition of the strain rate as the descent rate of the indenter divided by its instantaneous depth (\dot{h}/h) and second, considering the relationship between the strain rate and the load and hardness variations during the test. Based on the findings, the strain rate is greatly larger at shallow indentations, and the depth beyond which it attains the constant value depends on the material and the set \dot{P}/P ratio. Lastly, incorporating the relationship between the hardness and strain rate, it is revealed that although the strain rate variation changes the material hardness, its effect does not give a justification for the observed indentation size effect (ISE); therefore, other contributing parameters are discussed for their possible effects on this phenomenon.

Keywords

Glassy polymers · Amorphous · Nanoindentation · Hardness · Indentation strain rate · Continuous stiffness measurement · Loading rate · Poly(methyl methacrylate) · Polycarbonate · Elastic modulus · Polymeric glasses

Introduction

The interest of many researchers has been recently directed to study the mechanical properties of polymers in small size scales due to their extensive use in nano- and microscale elements over the past decades. A large series of constitutive models and experimental methods by which the properties of this class of materials can be thoroughly captured in millimeter or larger size scales has been presented (Hasan et al. 1993; Hoy and Robbins 2006; Van Breemen et al. 2012; Mulliken and Boyce 2006; Anand and Gurtin 2003; Voyiadjis and Samadi-Dooki 2016); however, there are still many questions about the behavior and deformation mechanism of polymers in submicron size scales. To acquire the precise and reliable results for mechanical properties, including elastic modulus and hardness, of very small volumes of materials, instrumented-indentation testing (IIT) can be employed (Al-Haik et al. 2004; Boersma et al. 2004; Lee et al. 2004; Zeng et al. 2012). In this technique, an indenter induces a localized deformation by applying a specified load on the material surface. Basically, there are two different indentation methods: (1) basic mode in which, with monotonic loading and unloading, the mechanical properties are only measured at the predefined maximum load from the unloading curve, and (2) continuous stiffness measurement (CSM) mode in which a small oscillation force is superimposed on the primary loading signal, and the resulting response of the system is analyzed through a frequency-specific amplifier. Employing the second procedure, the material mechanical properties can be continuously measured from zero to the maximum indentation depth during the loading segment. In addition,

while in the former, the contact stiffness is measured just at the initial point of the unloading, the measurement of the contact stiffness at any point along the loading segment is possible in the latter with a smaller time constant (Li and Bhushan 2002; Hay et al. 2010; Pethica and Oliver 1988). Therefore, the small time constant of the CSM method makes it more useful for measuring the properties of materials especially those which are strongly time dependent like polymers. The loading (or strain) rate is controlled in a different way in these two modes; during the basic mode, the load is applied with a constant rate on the sample surface by the indenter until it reaches a determined maximum value; however, in the CSM mode, the indenter travels up to a predefined maximum depth and the load is controlled so that the loading rate divided by the load (\dot{P}/P) remains constant over the course of a single indentation.

The loading and strain rates are adjustable parameters in nanoindentation experiments, and their variations have shown profound effects on the mechanical response of time-dependent materials like polymers (Odegard et al. 2005; Mazeran et al. 2012; Samadi-Dooki et al. 2016; Malekmotiei et al. 2015; Kraft et al. 2001; White et al. 2005; Zhang et al. 2009; Shen et al. 2004). That being the case, a closer look at the strain (loading) rate variation during the indentation is required since the generated strain and stress fields in the material due to the loading by a self-similar tip is inhomogeneous. In the basic mode nanoindentation, although the test is conducted with a constant loading rate \dot{P} , the strain rate is considerably decreasing at shallow depths, and it eventually approaches almost stable value after a long distance travel of the tip into the material. For this reason, an average value of the strain rate over the deep part of the indentation can be considered as the representative strain rate of the test (Schuh and Nieh 2003). On the other hand, in the CSM nanoindentation experiments, the \dot{P}/P ratio is set as a constant value at the beginning of the test. It has been shown that the indentation strain rate can also be assumed to remain constant during the constant \dot{P}/P experiment where the material hardness has the steady-state value, i.e., $\dot{H} = 0$ (Lucas and Oliver 1999). However, the indentation size effect (ISE), which is the increment of hardness as the indentation depth decreases, has been observed during the nanoindentation experiments on many materials including crystalline and amorphous solids (Briscoe et al. 1998; Voyiadjis and Zhang 2015). In a study on Al-based foams, it has been observed that the strain rate varies about three orders of magnitude during the first 200 nm of the indentation before reaching a steady-state value (Kraft et al. 2001). As a result, in the case of the CSM mode, the indentation strain rate can be considered constant in that part of the test where the ISE is negligible.

Conducting the CSM nanoindentation experiments on PMMA and PC as polymeric glasses, the variation of the strain rate during the course of a single test is investigated as a main goal in this chapter. Examining the variation of the \dot{P}/P ratio during the test shows that although the \dot{P}/P ratio is set to remain invariant during the loading segment, it takes a considerable tip travel distance until it stabilizes and reaches the set value. Furthermore, the indentation strain rate, which has been incorrectly considered as the \dot{P}/P ratio in some studies (Shen et al. 2004, 2006; Vachhani et al. 2013), is also found to change at shallow depths of

indentation. The obtained results show a good correlation between the instantaneous indentation strain rate, which is evaluated directly from the indentation depth-time data recorded during the loading segment, and the strain rate relation proposed by Lucas and Oliver (1999) based on the variation of the load and hardness. As another purpose, the possible relation between the variation of the strain rate during the nanoindentation and the observed ISE in polymers is also scrutinized in this chapter. While the high values of strain rate in shallow depths can cause the increment of material hardness, it is discussed here that it cannot be the reason for the observed ISE since the obtained high values of hardness could be the result of the indentation strain rates which are orders of magnitude higher than the actual recorded strain rate values.

Materials and Methods

Sample Preparation

The commercially manufactured (Goodfellow, Cambridge, UK) polymeric glasses including PMMA and PC, 2.0 and 5.0 mm-thick sheets, respectively, are considered for this investigation. The sheets are first cut into 20×20 mm squares, and then washed with 30% isopropyl alcohol (IPA) to eliminate the remainders of the protective film, and at the end rinsed with distilled water. Using a TA Instruments 2920 differential scanning calorimetry (DSC) machine, the glass transition temperature (T_g) of the specimens is measured to be about 110 °C and 148 °C for PMMA and PC, respectively. The samples are annealed at 120 °C for 4 h to remove any thermal history, and then cooled down to ambient temperature with the rate of 10 °C/h in a vacuum oven. The roughness of sample surface is one of the factors which affects the nanoindentation results since high values of roughness can make inaccuracy in the hardness of material measurements; therefore, to capture the surface topography of the samples, an Agilent 5500 atomic force microscope (AFM) is utilized. Since the average surface roughness, R_a , of the PC and PMMA specimens are 0.411 ± 0.033 and 0.372 ± 0.013 nm, respectively, one can assume the flat surface for samples (Kim et al. 2007), and there is no need to modify the obtained results for materials' hardness (Voyiadjis and Malekmotiei 2016).

Nanoindentation Analysis

To address the goal of this chapter which is scrutinizing the strain rate variation during the course of a single nanoindentation experiment and its effect on the observed ISE, an MTS Nanoindenter[®] XP equipped with a three-sided pyramidal Berkovich diamond tip is employed (Voyiadjis and Malekmotiei 2016). The mechanical properties of the specimens are measured through the CSM mode indentation in which the load-hold-unload sequences are carried out with the constant \dot{P}/P during the loading stage. According to the formulations developed by Oliver and Pharr

(2004), the material hardness is defined as the mean contact pressure under the indenter as follows:

$$H = \frac{P}{A_c} \quad (1)$$

where P is the applied load on the sample surface and A_c is the projected contact area of the hardness impression at that load. Based on this description, a precise measurement of the contact area between the sample surface and the indenter tip is required to calculate the hardness. The contact area is a function of the contact depth, h_c , and equal to $A_c = 24.56h_c^2$ for a perfect Berkovich indenter tip. The one used for these experiments is not ideally sharp; so, the contact area function is obtained through calibrating the tip which improves the accuracy of the contact area measurements by accounting for the tip imperfections and leads to introducing some additional terms to the above relation as:

$$A_c = 24.56h_c^2 + C_1h_c^1 + C_2h_c^{1/2} + C_3h_c^{1/4} + \dots + C_8h_c^{1/128} \quad (2)$$

in which C_1 through C_8 are constant coefficients which are obtained based on the results of the nanoindentation on fused silica as a standard sample. Another important parameter that needs to be accurately determined is the depth over which the material is in contact with the tip (h_c). The contact depth is estimated using

$$h_c = h - \varepsilon \frac{P}{S}. \quad (3)$$

where h is the total penetration depth, S is the elastic contact stiffness, and ε is a constant that depends on the indenter geometry (for a Berkovich indenter $\varepsilon = 0.75$ (Oliver and Pharr 1992)). As already stated, the CSM technique makes the continuous measurement of the contact stiffness as a function of depth possible during the loading segment of the indentation. Considering the imposed driving force as $P = P_0 e^{i\omega t}$ and the indenter displacement response as $h(\omega) = h_0 e^{i(\omega t + \alpha)}$, the elastic contact stiffness is calculated as follows:

$$S = \left[\frac{1}{\frac{P_0}{h(\omega)} \cos(\alpha) - (K_s - m\omega^2)} - \frac{1}{K_f} \right]^{-1} \quad (4)$$

in which P_0 is the force oscillation magnitude, ω is the oscillation frequency, h_0 is the resulting displacement oscillation magnitude, and α is the phase angle between the displacement and force signals. The other contributing parameters are the leaf spring constant, K_s , that supports the indenter, the indenter mass, m , and the indenter frame stiffness, K_f (Li and Bhushan 2002).

Another mechanical property measured in the nanoindentation experiments is the elastic modulus of the sample, E , which is calculated by the following relation:

$$\frac{1}{E_r} = \frac{1 - \nu^2}{E} + \frac{1 - \nu_i^2}{E_i}. \quad (5)$$

where E_r is the reduced elastic modulus which attributes to the elastic deformation in both the sample and indenter, ν is the sample Poisson's ratio, and E_i and ν_i are the indenter elastic modulus and Poisson's ratio, respectively. Sneddon (1965) has developed a relation for the reduced elastic modulus as follows:

$$E_r = \frac{S}{2\beta} \sqrt{\frac{\pi}{A_c}}. \quad (6)$$

where β as a constant depends on the indenter geometry and is about 1.034 for the Berkovich tip. Subsequently, as a main feature of the CSM method, the material hardness and elastic modulus are measured as continuous functions of depth with the course of an individual loading-unloading cycle.

The Indentation Strain Rate

Basically in the nanoindentation experiments, the strain rate affects the material in a direction perpendicular to the sample surface and is correlated with the displacement/loading rate of the indentation. For a pyramidal indenter, the indentation strain rate is defined as the penetration rate of the indenter into the material divided by its instantaneous depth as follows (Mayo and Nix 1988):

$$\dot{\epsilon}_i = \left(\frac{1}{h}\right) \left(\frac{dh}{dt}\right). \quad (7)$$

where t is time. In a study, Lucas and Oliver (1999) investigated that by keeping the loading rate divided by the load (\dot{P}/P) constant during the CSM nanoindentation, the indentation strain rate can also remain constant. It has been shown that incorporating the loading and hardness data, the indentation strain rate can be obtained as (Lucas and Oliver 1999):

$$\dot{\epsilon}_i = \frac{\dot{h}}{h} = \frac{1}{2} \left(\frac{\dot{P}}{P} - \frac{\dot{H}}{H} \right). \quad (8)$$

in which \dot{H} is the hardness variation rate and other parameters are defined before. According to Eq. 8, the indentation strain rate reaches a constant value $\left(\frac{1}{2} \frac{\dot{P}}{P}\right)$ at large indentation depths where the material hardness is almost unvaried, i.e., $\dot{H} = 0$.

It is noteworthy to mention that two main simplifying assumptions have made to get this relation: (a) the projected contact area relation is considered as $A = 24.56h^2$ which is used for an ideal Berkovich indenter tip, and (b) instead of the contact or plastic depth, the total depth is used in the contact area function.

Experimental Procedure

As mentioned in preceding sections, as a first step, the specimens should be thoroughly prepared and the tip should be carefully calibrated; the tests then are triggered by running the loading-hold-unloading cycles as follows: before any measurement, the tip drift should be controlled at a rate below 0.05 nms^{-1} , so it is held on the top of the sample surface until it gains the stabilized rate. The tip then moves downward to reach the material surface. As soon as the tip touches the sample surface, the loading stage begins with a constant \dot{P}/P ratio and it continues until a specified maximum depth of $10 \text{ }\mu\text{m}$. To account for the creep behavior of the polymer, the load is then held at this stage for 10 s , and eventually, the unloading part is carried out with a constant unloading rate until 10% of the maximum load. Since the goals are investigating the variation of \dot{P}/P ratio during the whole nanoindentation experiment from zero to the maximum depth, and also its contribution on the indentation strain rate, a series of tests are performed on annealed PMMA and PC samples with three different set values of \dot{P}/P (0.005 , 0.05 , and 0.11 s^{-1}). For each \dot{P}/P ratio, 25 indents are accomplished to get the accurate results and to prevent from interaction of the indents, $150 \text{ }\mu\text{m}$ distance is considered between them.

Results and Discussion

Variation of the \dot{P}/P in the Course of an Indentation

The applied load on the sample and the tip travel distance are recorded as unbroken curves in the CSM nanoindentation experiments with nN and sub- nm exactness, respectively. Since the loading rate divided by the load is constant during the loading stage, i.e., $\dot{P}/P = \eta$, the load is expected to be an exponential function of time as follows:

$$P = \beta e^{\eta t}. \quad (9)$$

where β is the constant obtained by solving the ordinary differential equation (ODE). The load variation with time during the loading section of the nanoindentation on PC sample is presented in Fig. 1 (Voyiadjis and Malekmotiei 2016); the figure shows the results for three different set \dot{P}/P ratios and their exponential interpolations (lines) for comparison. The result curves depart from the exponential behavior at shallow indentation depths while they behave in accordance with the exponential variation within the long tip travel distance, especially for higher \dot{P}/P values. This discrepancy at the early stages of loading is due to the fact that based on Eq. 9, the initial loading condition is $P(0) = \beta$; however, the set initial condition for the experiment process is $P(0) = 0$. Thus, the indenter \dot{P}/P ratio can be adjusted to the set \dot{P}/P value after several nanometers of indentation displacement (or several seconds).

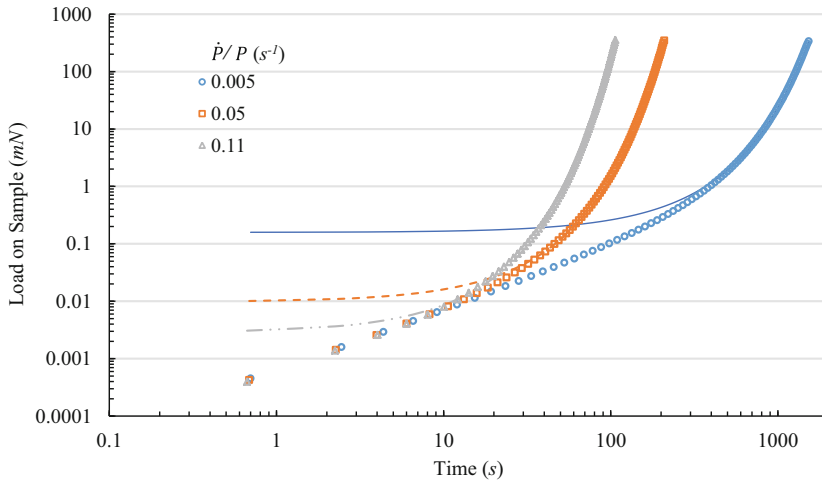


Fig. 1 Variation of load versus time response of the indenter for PC sample measured on the loading segment of the nanoindentation experiment at three different set \dot{P}/P values. The *lines* represent the exponential interpolations (Reprinted from Voyiadjis and Malekmtotiei 2016)

The actual variation of \dot{P}/P ratio as a function of the indentation depth is presented in Fig. 2 for PC and PMMA (Voyiadjis and Malekmtotiei 2016) at three different set values of this ratio which are shown as horizontal dashed lines. As mentioned before, the actual \dot{P}/P values do not approach their set values right after the indenter tip comes into the contact with the sample surface. Furthermore, the depth at which the \dot{P}/P ratio stabilizes does not depend on the material; however, it is extremely dependent on the set \dot{P}/P value. As depicted in Fig. 2 (Voyiadjis and Malekmtotiei 2016), the depth beyond which the actual \dot{P}/P approaches the set value and stabilizes is smaller for the bigger set value of \dot{P}/P : it is almost 1000 nm for the set \dot{P}/P value of 0.005 s^{-1} and reduces to 200 and 100 nm for the set \dot{P}/P values of 0.05 and 0.11 s^{-1} , respectively. Interestingly, the starting point of actual value of \dot{P}/P is not dependent on the material and set \dot{P}/P ratio, and it approximately equals 0.3 s^{-1} for all experiments.

Variation of the Strain Rate in the Course of an Indentation

In the CSM nanoindentation method, since the tip displacement is recorded continuously with time, the indentation strain rate can be directly calculated by using Eq. 7 and simple numerical differentiation as a continuous function of the indentation depth. In addition, indirect evaluation of the indentation strain rate during the loading segment of the test is possible by incorporating Eq. 7 and using the recorded load on the sample and the measured material hardness as functions of the depth. Since to employ Eq. 8 the variation of the instantaneous hardness rate divided by hardness (\dot{H}/H) is required, this parameter is represented in Fig. 3 (Voyiadjis and Malekmtotiei 2016) at three different set \dot{P}/P ratios for PC and PMMA. As depicted

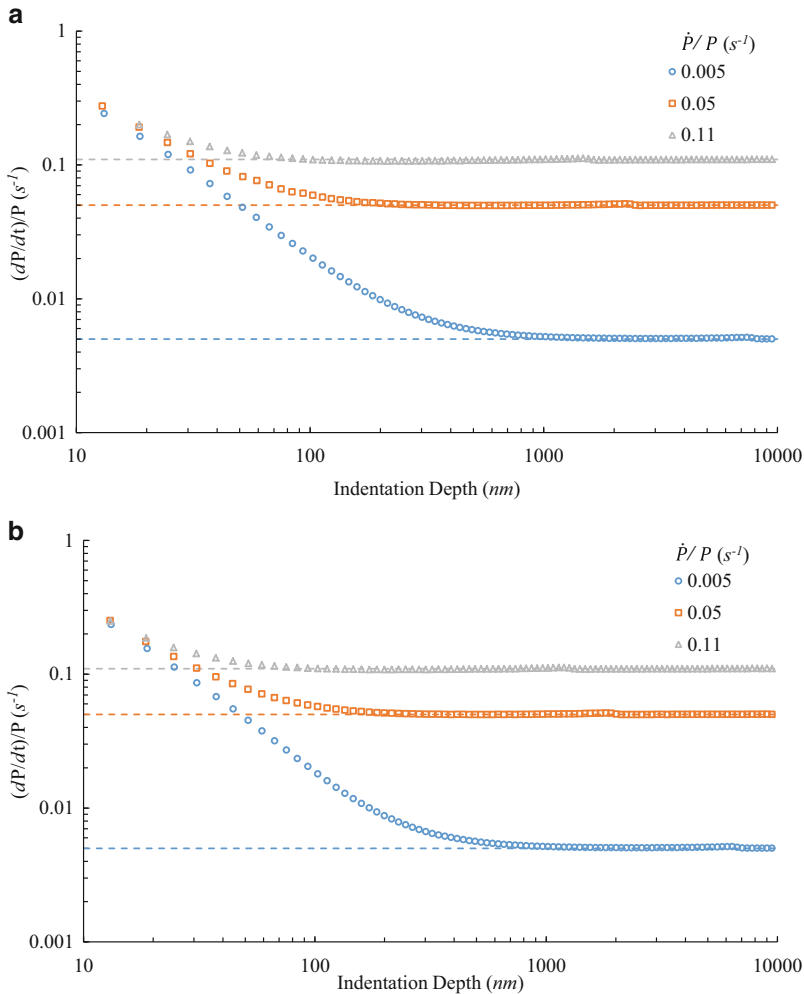


Fig. 2 Variation of the actual \dot{P}/P values with indentation depth at three different set \dot{P}/P values for (a) PC and (b) PMMA (Reprinted from Voyiadjis and Malekmotiei (2016))

in Fig. 5, the material hardness is higher at shallower indentation depths (indentation size effect) and then reaches a plateau at the certain depth which is the representative of the macroscopic hardness. Therefore, this trend results in the negative values of the \dot{H}/H ratio at the initial stages of the loading section and finally zero values of \dot{H}/H at deep part of the indentation (see Fig. 3). As another result obtained from Fig. 3, the depth beyond which \dot{H} can be assumed zero depends on the material and the set \dot{P}/P value.

Figure 4 (Voyiadjis and Malekmotiei 2016) displays the variation of the indentation strain rate with the indentation depth calculated based on the two different

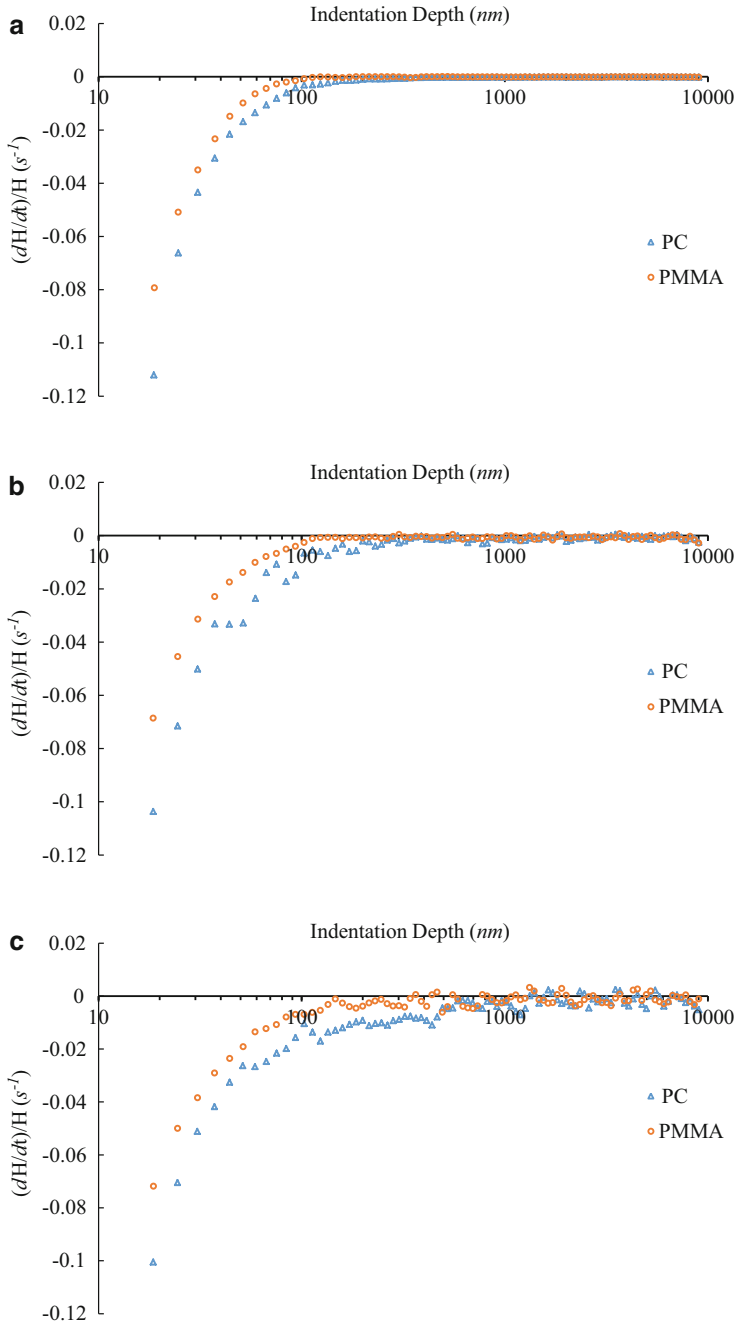


Fig. 3 Variation of the instantaneous hardness rate divided by hardness $\left(\frac{\dot{H}}{H}\right)$ with indentation depth for PC and PMMA at set \dot{P}/P equal to (a) 0.005, (b) 0.05, and (c) 0.11 s^{-1} (Reprinted from Voyiadjis and Malekmoie (2016))

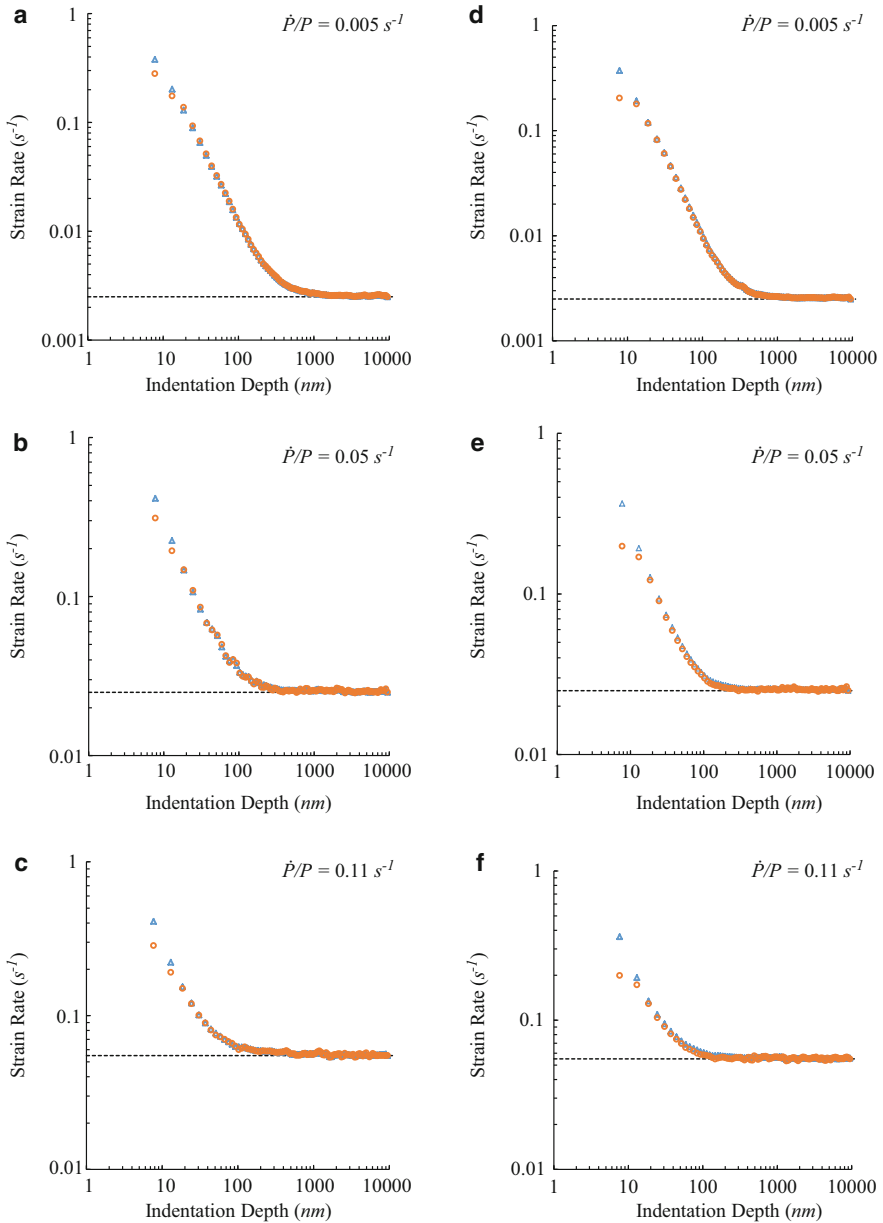


Fig. 4 Variation of the indentation strain rate based on Eq. 7 (Δ) and Eq. 8 (\circ) versus indentation depth for experiments on PC (**a–c**) and PMMA (**d–f**) at three different set \dot{P}/P values. The dashed lines represent the expected values of strain rate $\left(\frac{1}{2} \frac{\dot{P}}{P}\right)$ (Reprinted from Voyiadjis and Malekmoetie (2016))

approaches. Since both methods show almost the same results, it reveals the validity of the assumptions made by Lucas and Oliver (1999) to obtain Eq. 8 not only for deep but also shallow indentations of polymeric glasses. The most important result is that the indentation strain rate is not constant during the loading segment of the CSM nanoindentation of glassy polymers, and its variation is material and rate dependent. However, it can be assumed to be constant and equal to $\frac{1}{2} \frac{\dot{P}}{P}$ for deep enough indentation experiments.

Indentation Size Effect

In rate-dependent materials including polymers, the flow stress extremely depends on the applied loading (strain) rate: the higher the experiment strain rate, the higher the yield stress (Voyiadjis and Samadi-Dooki 2016; Samadi-Dooki et al. 2016; Malekmtiei et al. 2015; Richeton et al. 2006; Rottler and Robbins 2003). Therefore, since there is a relation between the flow stress and hardness of the material through Tabor's relation (Prasad et al. 2009), the higher value of hardness is expected from nanoindentation with the higher strain rate. The variation of hardness versus the tip displacement is presented in Fig. 5 (Voyiadjis and Malekmtiei 2016) for experiments on PC and PMMA samples at three different set \dot{P}/P values. It is observed that the obtained hardness values are higher as the strain rate increases; especially, the macroscopic hardness which is the hardness at the deep part of the nanoindentation and is the plateau for each curve depends on both material and strain rate. It is clear in these figures that the strain rate dependency of PMMA is more considerable which is, physically, demonstrated as smaller shear activation volumes in this material (Malekmtiei et al. 2015). Another observation in Fig. 5 is the profound increment of the hardness as the depth decreases during each indentation which is known as the ISE at nanoscales (Shen et al. 2006; Lam and Chong 1999; Zhang and Xu 2002). Due to the above-mentioned reason, the increased values of the material hardness at shallow indentation depths might be correlated with the higher values of the strain rate at these depths. However, an exact quantitative analysis is needed to understand and evaluate this possible relationship.

Many studies show that there is a linear relationship between the flow stress (or hardness) of the polymeric glasses and the logarithm of the strain rate. The explicit relationships have been previously obtained for PC and PMMA by nanoindentation evaluations (Samadi-Dooki et al. 2016; Malekmtiei et al. 2015). Using the obtained formulations and the strain rate variation during indentation, the fictitious hardness can be calculated for each test. It should be mentioned that to calculate the fictitious hardness, it is assumed that the hardness variation is just the result of the strain rate variation during the loading stage. Figure 6 shows the actual measured hardness as well as the calculated hardness versus the indentation depth for PC and PMMA at three different set \dot{P}/P values for comparison. The curves of PC sample show that the calculated hardness is almost constant and there is no considerable change

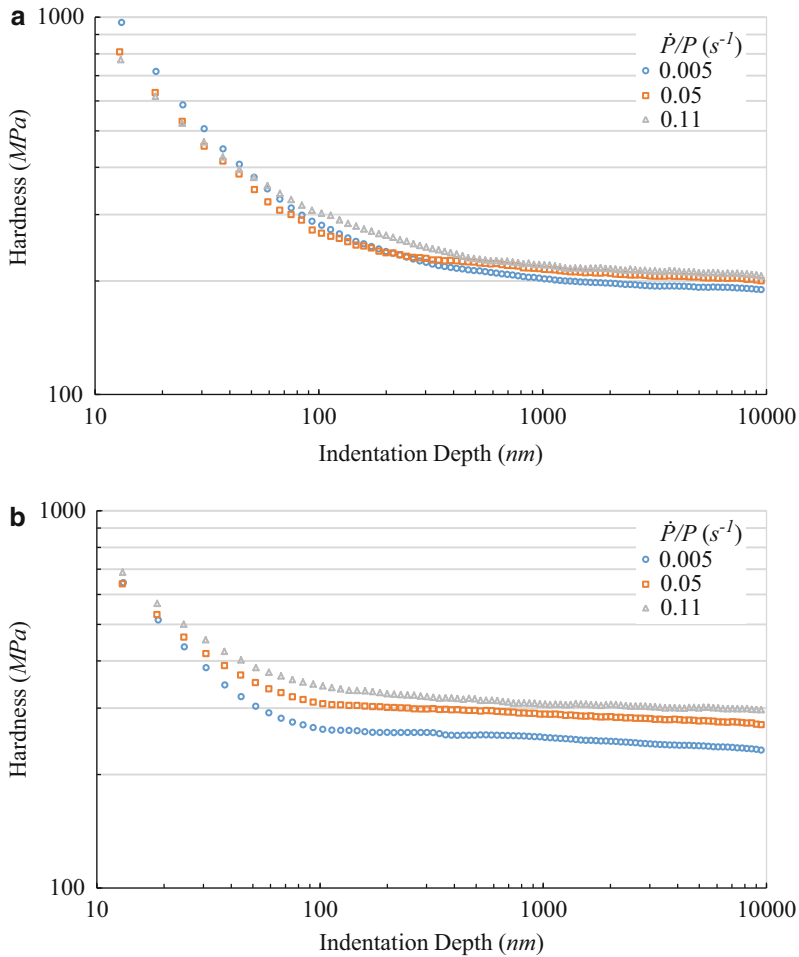


Fig. 5 Variation of the hardness versus indentation depth during the loading segment of the experiments on (a) PC and (b) PMMA at three different set \dot{P}/P values (Reprinted from Voyiadjis and Malekmoitei (2016))

during a test; however, its variation during each indentation on PMMA sample is notably large in amount and follows almost the same hardening pattern at shallow depths as the actual hardness variation trend. In addition, for PMMA, the calculated hardness is the same (about 336 MPa) for all \dot{P}/P values when the loading stage is triggered, which is acceptable for the reason that the strain rate at the beginning of the indentation is also the same for different \dot{P}/P ratios (Fig. 4), while the actual hardness at this point (maximum hardness in each curve) is different for different \dot{P}/P values in both PC and PMMA. More importantly, as Fig. 4 represents the strain rate variation in the course of an indentation is approximately material

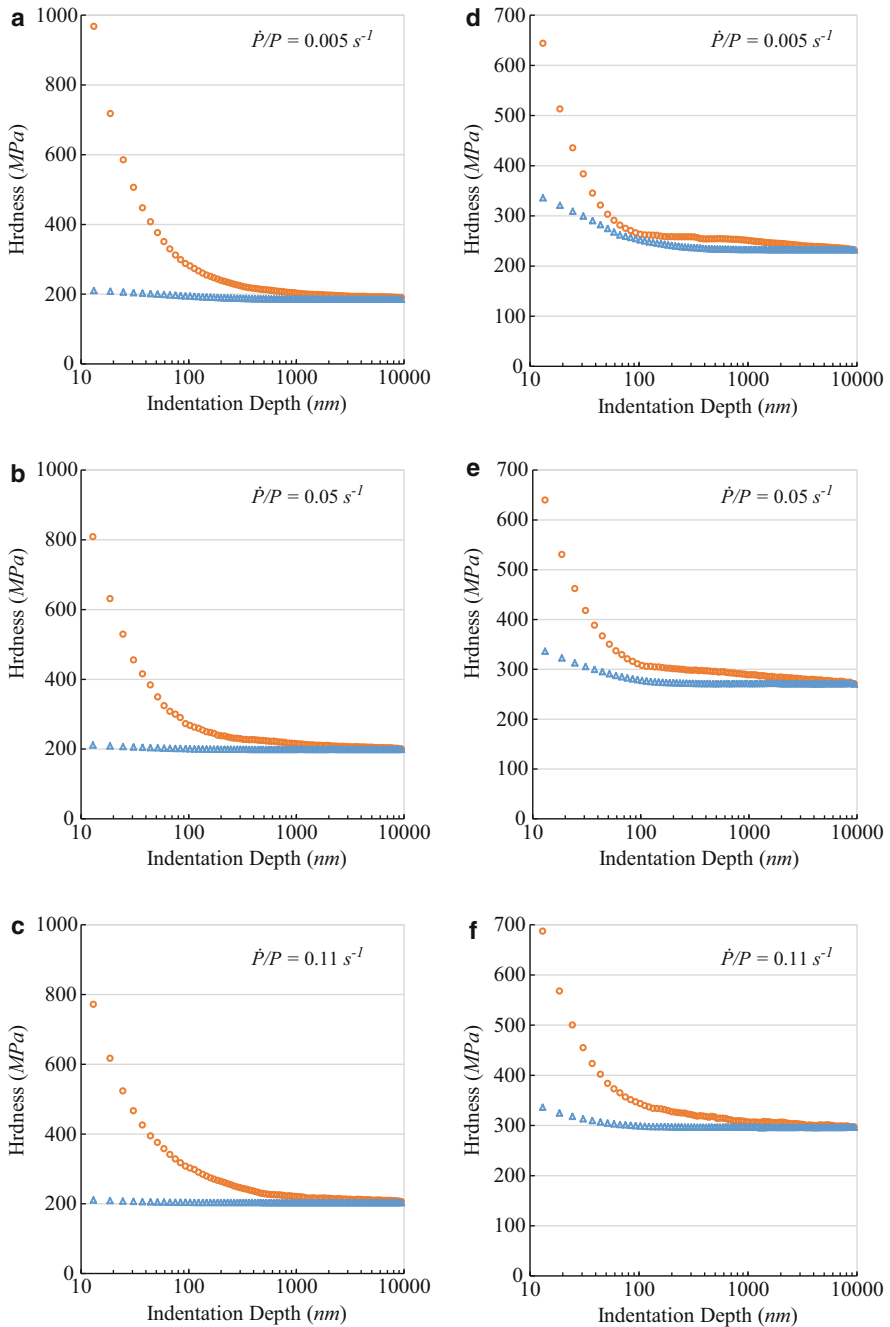


Fig. 6 Variation of the calculated (Δ) and actual (\circ) hardness versus indentation depth during the loading segment of the experiments on PC (a–c) and PMMA (d–f) at three different set \dot{P}/P values (Reprinted from Voyiadjis and Malekmoitei (2016))

independent, however, PMMA reveals a more intense strain rate dependency of hardness response than PC (Fig. 5); therefore, a more profound indentation size effect is expected to be observed in the nanoindentation of PMMA. Nevertheless, the real situation is different since the hardness variation at shallow depths is more noticeable for PC in comparison with PMMA (see Figs. 5 and 6) (Voyiadjis and Malekmtotiei 2016). The main result from these observations is that although the strain rate variation during the loading segment of the CSM nanoindentation on PC and PMMA is notable, it cannot be the major cause of the observed ISE in amorphous polymers. As a matter of fact, the indentation strain rate variation during the loading has *no* contribution to the observed ISE of PC and its contribution to the ISE phenomenon in PMMA is *negligible*. Additionally, assuming the constant strain rate during the CSM nanoindentation of polymeric glasses for the size effect studies seems to be reasonable and there should exist other mechanisms behind this phenomenon which are correlated to the localization and or free surface effects (Alisafaei and Han 2015; Han et al. 2016).

It is noteworthy to mention that another important factor which can affect the contact area and, subsequently, the measured hardness, especially at shallow depths of the indentation, is the material pile-up around the indenter tip. In Fig. 6, since the calculated hardness is obtained from the direct measurement of the indentation strain rate (\dot{h}/h) (Samadi-Dooki et al. 2016; Malekmtotiei et al. 2015), it is not affected by the material pile-up, while the actual measured hardness in this figure could be affected by the pile-up. For this reason, the material pile-up around the tip could be another factor that causes the difference between the calculated hardness of material and the actual one.

Another important phenomenon which is usually observed during the CSM nanoindentation experiments is a small size effect on the recorded elastic modulus of the material. As shown in Fig. 7 (Voyiadjis and Malekmtotiei 2016), it is an increased Young's modulus at shallower indentation depths. This phenomenon is in contrast to the earlier observations from particle embedment experiments (Teichroeb and Forrest 2003; Karim and McKenna 2011, 2012, 2013; Hutcheson and McKenna 2007). As discussed comprehensively in the literature (Parry and Tabor 1973, 1974) the applied hydrostatic pressure on the polymer samples can hamper the chain movements which are required for relaxation processes and can subsequently result in a considerable increment of the glass transition temperatures of the material. Based on that, it has been proposed that, in nanoindentation experiments, the contact loading at the indenter tip-polymer interface induces hydrostatic pressure under the tip which increases the glass transition temperature of the sample near the surface, and correspondingly, the increased stiffness of the material at low indentation depths has been related to the increment of T_g (Gacoin et al. 2006; Tweedie et al. 2007). Therefore, the observed considerable material stiffening at shallow indentation depths (for depths of <50 nm in Fig. 7) (Voyiadjis and Malekmtotiei 2016) could also attribute to the elevated values of T_g at the surface layer within this tip travel distance compared to the bulk. Moreover, incorporating the shear transformation theory, Voyiadjis and Samadi-Dooki (2016) have proposed a model for yielding

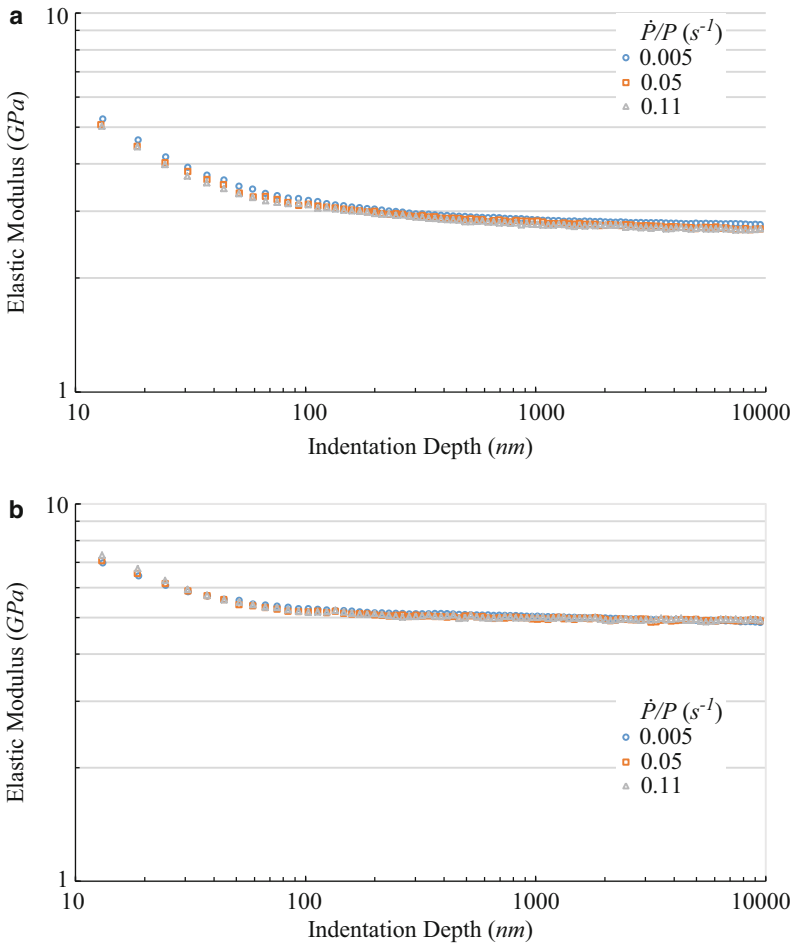


Fig. 7 Variation of the elastic modulus versus indentation depth during the loading segment of the experiments on (a) PC, and (b) PMMA at three different set \dot{P}/P values (Reprinted from Voyiadjis and Malekmoitei (2016))

and plasticity of amorphous polymers which shows the relationship between the yield stress (which is proportional to hardness through Tabor's relation) and an activation energy which itself is a function of the elastic modulus of the material. Thus, as another confirmation, there exists a possibility that the increased hardness observed at the shallow indentation depths to be interrelated to the increased elastic modulus at these depths. This hypothesis, however, should be viewed as a qualitative observation and treated cautiously since there is a noticeable difference between the length scales during which the elastic modulus and hardness increments are observed (see Figs. 5 and 7).

Conclusions

The mechanical behavior of viscoelastic-viscoplastic materials, including polymers, strongly depends on the rate at which they are loaded. In the case of the nanoindentation, where these behaviors are evaluated at nanoscales, there are different methods in which the loading rate can be controlled in a way that the strain rate changes or remains constant during the experiment. For many reasons, for example, to study the effect of the temperature, thermal history, composition of alloys, etc., it is desirable to conduct a test during which the strain rate remains constant. The continuous stiffness measurement (CSM) nanoindentation is a technique which offers testing at constant loading rate to load ratio (\dot{P}/P) ratios; this has been manifested as constant strain rates during the test. In this chapter, the assumption of the constant strain rate during this nanoindentation technique is studied for glassy polymers. Investigating the instantaneous variations of \dot{P}/P and strain rate during the CSM nanoindentation on poly(methyl methacrylate) (PMMA) and polycarbonate (PC) samples showed that this assumption is not valid during the whole test and the strain rate changes in the early stages of the indentation before acceptable stabilization; the depth beyond that this parameter stabilizes depends on the material and the set \dot{P}/P value. It has been shown that although by assuming the constant value of \dot{P}/P , an exponential load-time response of the indenter is expected, the load-time curves do not obey the exponential variation in early stages of the loading. The reason is the incompatibility of the actual initial load with the initial condition of the exponential loading which is required to assure a constant \dot{P}/P . To overcome this problem, one can apply a very small load prior to the main loading segment of the experiment; this adjusts the aforementioned incompatibility of the initial condition. By this adjustment, it is expected that the \dot{P}/P value during the whole test remains constant and equals the initial set value. However, it may not lead to a constant strain rate since the hardness variation rate also contributes to the strain rate variation as proposed by Lucas and Oliver (1999). The strain rate can be assumed constant only at deep enough indentations where the load-displacement curve obeys the Hertzian relation (Malekmotiei et al. 2015). In this chapter, the possible effect of the variation of the strain rate during the indentation on the observed indentation size effect (ISE) in amorphous polymers has also been discussed (Voyiadjis and Malekmotiei 2016). While it is concluded that the increased strain rate within the shallow indentation depths cannot be the sole reason for the observed profound ISE, contribution of the other factors, such as material pile-up around the tip and stiffening due to the glass transition temperature (T_g) shift induced by the hydrostatic component of the stress, have been qualitatively discussed.

References

- M. Al-Haik, H. Garmestani, D. Li, M. Hussaini, S. Sablin, R. Tannenbaum, K. Dahmen, *J. Polym. Sci. B Polym. Phys.* **42**, 1586 (2004)
F. Alisafaei, C.-S. Han, *Adv. Condens. Matter Phys.* **2015**, 391579 (2015)

- L. Anand, M.E. Gurtin, *Int. J. Solids Struct.* **40**, 1465 (2003)
- A. Boersma, V. Soloukhin, J. Brokken-Zijp, G. De With, *J. Polym. Sci. B Polym. Phys.* **42**, 1628 (2004)
- B. Briscoe, L. Fiori, E. Pelillo, *J. Phys. D. Appl. Phys.* **31**, 2395 (1998)
- E. Gacoin, C. Fretigny, A. Chateauinois, A. Perriot, E. Barthel, *Tribol. Lett.* **21**, 245 (2006)
- C.-S. Han, S.H. Sanei, F. Alisafaei, *J. Polym. Eng.* **36**, 103 (2016)
- O. Hasan, M. Boyce, X. Li, S. Berko, *J. Polym. Sci. B Polym. Phys.* **31**, 185 (1993)
- J. Hay, P. Agee, E. Herbert, *Exp. Tech.* **34**, 86 (2010)
- R.S. Hoy, M.O. Robbins, *J. Polym. Sci. B Polym. Phys.* **44**, 3487 (2006)
- S. Hutcheson, G. McKenna, *Eur. Phys. J. E.* **22**, 281 (2007)
- T.B. Karim, G.B. McKenna, *Polymer* **52**, 6134 (2011)
- T.B. Karim, G.B. McKenna, *Macromolecules* **45**, 9697 (2012)
- T.B. Karim, G.B. McKenna, *Polymer* **54**, 5928 (2013)
- J.Y. Kim, S.-K. Kang, J.-J. Lee, J.-i. Jang, Y.-H. Lee, D. Kwon, *Acta Mater.* **55**, 3555 (2007)
- O. Kraft, D. Saxa, M. Haag, A. Wanner, *Z. Metallkd.* **92**, 1068 (2001)
- D.C. Lam, A.C. Chong, *J. Mater. Res.* **14**, 3784 (1999)
- C. Lee, J. Kwon, S. Park, S. Sundar, B. Min, H. Han, *J. Polym. Sci. B Polym. Phys.* **42**, 861 (2004)
- X. Li, B. Bhushan, *Mater. Charact.* **48**, 11 (2002)
- B. Lucas, W. Oliver, *Metall. Mater. Trans. A* **30**, 601 (1999)
- L. Maleknotiei, A. Samadi-Dooki, G.Z. Voyiadjis, *Macromolecules* **48**, 5348 (2015)
- M. Mayo, W. Nix, *Acta Metall.* **36**, 2183 (1988)
- P.E. Mazeran, M. Beyaoui, M. Bigerelle, M. Guigon, *Int. J. Mater. Res.* **103**, 715 (2012)
- A. Mulliken, M. Boyce, *Int. J. Solids Struct.* **43**, 1331 (2006)
- G. Odegard, T. Gates, H. Herring, *Exp. Mech.* **45**, 130 (2005)
- W.C. Oliver, G.M. Pharr, *J. Mater. Res.* **7**, 1564 (1992)
- W.C. Oliver, G.M. Pharr, *J. Mater. Res.* **19**, 3 (2004)
- E.J. Parry, D. Tabor, *J. Mater. Sci.* **8**, 1510 (1973)
- E.J. Parry, D. Tabor, *J. Mater. Sci.* **9**, 289 (1974)
- J.B. Pethica, W.C. Oliver, *MRS Online Proc. Lib. Arch.* **130**, 13 (1988)
- K.E. Prasad, V. Keryvin, U. Ramamurty, *J. Mater. Res.* **24**, 890 (2009)
- J. Richeton, S. Ahzi, K. Vecchio, F. Jiang, R. Adharapurapu, *Int. J. Solids Struct.* **43**, 2318 (2006)
- J. Rottler, M.O. Robbins, *Phys. Rev. E* **68**, 011507 (2003)
- A. Samadi-Dooki, L. Maleknotiei, G.Z. Voyiadjis, *Polymer* **82**, 238 (2016)
- C.A. Schuh, T. Nieh, *Acta Mater.* **51**, 87 (2003)
- L. Shen, I.Y. Phang, T. Liu, K. Zeng, *Polymer* **45**, 8221 (2004)
- L. Shen, I.Y. Phang, T. Liu, *Polym. Test.* **25**, 249 (2006)
- I.N. Sneddon, *Int. J. Eng. Sci.* **3**, 47 (1965)
- J. Teichroeb, J. Forrest, *Phys. Rev. Lett.* **91**, 016104 (2003)
- C.A. Tweedie, G. Constantinides, K.E. Lehman, D.J. Brill, G.S. Blackman, K.J. Van Vliet, *Adv. Mater.* **19**, 2540 (2007)
- S. Vachhani, R. Doherty, S. Kalidindi, *Acta Mater.* **61**, 3744 (2013)
- L.C. Van Breemen, T.A. Engels, E.T. Klompen, D.J. Senden, L.E. Govaert, *J. Polym. Sci. B Polym. Phys.* **50**, 1757 (2012)
- G.Z. Voyiadjis, L. Maleknotiei, *J. Polym. Sci. Part B: Polym. Phys.* **54**, 2179 (2016)
- G.Z. Voyiadjis, A. Samadi-Dooki, *J. Appl. Phys.* **119**, 225104 (2016)
- G.Z. Voyiadjis, C. Zhang, *Mater. Sci. Eng. A* **621**, 218 (2015)
- C. White, M. Vanlandingham, P. Drzal, N.K. Chang, S.H. Chang, *J. Polym. Sci. B Polym. Phys.* **43**, 1812 (2005)
- F. Zeng, Y. Liu, Y. Sun, E. Hu, Y. Zhou, *J. Polym. Sci. B Polym. Phys.* **50**, 1597 (2012)
- T.-Y. Zhang, W.-H. Xu, *J. Mater. Res.* **17**, 1715 (2002)
- Y.F. Zhang, S.L. Bai, X.K. Li, Z. Zhang, *J. Polym. Sci. B Polym. Phys.* **47**, 1030 (2009)



Shear Transformation Zones in Amorphous Polymers: Geometrical and Micromechanical Properties **10**

George Z. Voyiadjis, Leila Malekmoitei, and Aref Samadi-Dooki

Contents

Introduction	334
Experimental Procedure	336
Sample Preparation	336
Nanoindentation Technique	337
Room-Temperature CSM Nanoindentation	339
High-Temperature Nanoindentation Procedure	340
Theory of Homogeneous Flow for Glassy Polymers	341
Results and Discussion	346
Calibrating the Nanoindentation Results	346
Shear Activation Volume	348
STZ's Activation Energy	352
STZ's Geometry	355
Concluding Remarks	356
References	357

Abstract

Glassy polymers are extensively used as high impact resistant, low density, and clear materials in industries. Due to the lack of the long-range order in the microstructures of glassy solids, plastic deformation is different from that in crystalline solids. Shear transformation zones (STZs) are believed to be the plasticity

G. Z. Voyiadjis (✉) · L. Malekmoitei
Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge,
LA, USA
e-mail: voyiadjis@eng.lsu.edu; lmalek1@lsu.edu

A. Samadi-Dooki
Computational Solid Mechanics Laboratory, Department of Civil and Environmental
Engineering, Louisiana State University, Baton Rouge, LA, USA
e-mail: asamad3@lsu.edu

carriers in amorphous solids and defined as the localized atomic or molecular deformation patches induced by shear. Despite a great effort in characterizing these local disturbance regions in metallic glasses (MGs), there are still many unknowns relating to the microstructural and micromechanical characteristics of STZs in glassy polymers. This chapter is aimed at investigating the flow phenomenon in polycarbonate (PC) and poly(methyl methacrylate) (PMMA) as glassy polymers and obtaining the mechanical and geometrical characteristics of their STZs. To achieve this goal, the nanoindentation experiments are performed on samples with two different thermal histories: as-cast and annealed, and temperature and strain rate dependency of the yield stress of PC and PMMA are studied. Based on the experimental results, it is showed that the flow in PC and PMMA is a homogeneous phenomenon at tested temperatures and strain rates. The homogeneous flow theory is then applied to analyze the STZs quantitatively. The achieved results are discussed for their possible uniqueness or applicability to all glassy polymers in the context of amorphous plasticity.

Keywords

Glassy polymers · Shear transformation zone · Nucleation energy · Shear activation volume · Homogeneous flow · Amorphous · Transformation shear strain · β -transition · Nanoindentation · Hardness · Plasticity

Introduction

Over the past several years, many researches have been conducted on investigating the mechanical behavior of polymers, due to the extensive use of them for design and development of a variety of components and structures. In general, polymers are divided in two different categories: semicrystalline and amorphous (glassy) polymers. Amorphous polymers are composed of entangled and disordered long molecular chains, and there is no significant chain alignment in their intra- and intermolecular structures. Since the molecular structure of the polymeric glasses (PGs) is totally different from that of the crystalline solids, the plastic deformation process does not obey the crystal plasticity rules. Moreover, in contrast to many crystalline solids, the postyield behavior begins with a softening at the onset of the yielding in the stress-strain characteristic behavior of PGs, and then continues by a hardening which starts at the specific strain and ends up at the break point (Boyce et al. 1988; Stoclet et al. 2010). There is a large body of literature dealing with the process of yielding and characterizing the mechanism of nonlinear elastoplastic deformation in glassy polymers, which resulted in different physical and phenomenological models (Ree and Eyring 1955; Robertson 1966; Argon 1973; Boyce et al. 1988; Arruda et al. 1995; Anand and Gurtin 2003; Mulliken and Boyce 2006; Chen and Schweizer 2011; Voyiadjis and Samadi-Dooki 2016).

In the context of crystal plasticity, crystal dislocations are the principal carriers of plasticity, and their slips result in plastic deformation (Argon 2008). But, since there is no long-range coherence in atomic or molecular structure of glassy solids, there are no analogous mobile defects. Consequently, the flow and the mechanism

of plastic response in the microstructural level are different from the crystalline solids. The ongoing widely recognized mechanism for the plastic response of all types of disordered solids, including metallic glasses, glassy polymers, and covalent glasses, is the cooperative localized rearrangement of molecular or atomic patches in small distinct regions which are called shear transformation zones (STZs) (Spaepen 1977; Falk 2007; Schuh et al. 2007; Pauly et al. 2010; Argon 2013). Especially, the presence of separate plastic deformation units has been experimentally recognized in glassy polymers and attributed to the formation of STZs (Oleinik et al. 2007).

The STZs are isolated irreversible stress relaxation events which form around free volume sites with the thermal fluctuations assistance under the action of an applied shear stress (Argon 1979; Falk et al. 2005; Argon 2013). The absence of long-range coherence in glassy materials results in the sessile transformations which, once formed, do not expand by translational movements of their interfaces. Therefore, the plasticity mediated by the shear transformations is nucleation controlled (Argon 1993; Spathis and Kontou 2001). Since PGs are considered as homogeneous and isotropic materials, a localized disturbance in their bulk can be considered as an Eshelby inclusion problem (Eshelby 1957). The Eshelby leading-edge homogenization method, which has been extensively applied for solving a broad area of problems in inhomogeneous media (Tandon and Weng 1984; Shodja et al. 2003; Malekmoie et al. 2013), is based on the strain compatibility of a medium containing no-elastic strains (Mura 1987). Using this method for amorphous solids, the STZ's nucleation energy, which is necessary to relate the shear flow stress to the shear flow strain rate through an Arrhenius function, can be obtained (Argon 2013). Based on the Eshelby solution, the microgeometrical and micromechanical properties of the embedded STZ, i.e., size, shape, and transformation shear strain, are determining parameters for evaluating the STZ's nucleation energy in glassy solids. Therefore, these characteristics have been extensively studied, especially for metallic glasses (Yang et al. 2007; Pan et al. 2008; Pan et al. 2009; Ju et al. 2011). It has been found that a single STZ in MGs possesses an average volume of less than 10 nm^3 including ≤ 500 atoms (Pan et al. 2008), with the average nucleation energy of about 1.5 eV (Yu et al. 2010) and transformation shear strain of 0.07 (Argon 2013).

In contrast, the number of studies devoted to evaluate the STZs quantitatively in glassy polymers is limited. The plastic deformation units pertaining parameters of some glassy polymers have been obtained by Argon and Bessonov (1977) based on double kink theory (Argon 1973). Later, in a molecular dynamics simulations, Mott et al. (1993) scrutinized the plastic deformation kinematics in amorphous atactic polypropylene, and found the transformation shear strain of about 0.015 in the spherical plastic flow units with the average 10 nm diameter. Moreover, Ho et al. (2003) also inquired the correlation between the STZ size scale and entanglement density by performing different compressive tests on mixable polystyrene-poly(2,6-dimethyl-1,4-phenylene oxide) (PS-PPO) blends at different mix ratios.

In this chapter, the nanoindentation technique is used to probe the flow nature of poly(methyl methacrylate) (PMMA) and polycarbonate (PC) as glassy polymers, and, consequently, their STZs' micromechanical and geometrical characteristics are

obtained. This chapter is organized as follows: in the first section, the experimental procedure is explained in two parts: sample preparation and the nanoindentation technique. In the second section, it is first shown that at tested temperatures and strain rates, the flow is homogeneous in these polymers. The homogeneous flow theory is then elaborated based on the Eshelby solution for nonelastic strain in an embedded inclusion in the representative volume element (RVE) and the Arrhenius function relating the shear flow stress to the shear flow strain rate. In the third section, the obtained results from the nanoindentation experiments on samples with two different thermal histories, as-cast and annealed, are presented. The characteristic properties of STZs including the nucleation energy barrier, size and shape of an STZ, and the shear activation volume are then obtained by utilizing the flow theory. Furthermore, the observed jump in the activation energy of the STZ is ascribed to the β -transition, and the energy barrier for this transition is found to be about 10% of the STZ nucleation energy. At the end, the concluding remarks are summarized in the last section.

Experimental Procedure

Sample Preparation

Commercially available 2.0- and 5.0-mm-thick sheets of amorphous poly (methyl methacrylate) (PMMA) and polycarbonate (PC), Goodfellow® catalogue #ME303020 and #L5433027, Cambridge, UK, are selected for this study, respectively. There is no preexisting molecular chain orientation in the sheets since they have been produced through the traditional method of cell cast. All the sheets are cut into $20 \times 20 \text{ mm}^2$ samples small enough for handling the nanoindentation experiments. The residues of the protective film covering the sheets are then removed by washing them with 30% isopropyl alcohol (IPA), and thoroughly rinsing by distilled water. To eliminate any moisture caused by the washing process, all the specimens are stored in a desiccator for at least 10 days before any experiment. The glass transition temperature (T_g) of the samples are measured by means of a TA Instruments 2920 differential scanning calorimetry (DSC) device, which is operating under the nitrogen flow and by using standard aluminum pans. The DSC cycles are conducted at $10 \text{ }^\circ\text{C min}^{-1}$ from ambient temperature to 200 and 250 $^\circ\text{C}$ for PMMA and PC, respectively. The calorimetric measurements reveal that the glass transition temperature of the PMMA and PC specimens are about 110 and 148 $^\circ\text{C}$, respectively.

Half of the samples, prepared as described above, are subjected to the following thermal treatment to study the effect of thermal history on their STZs' micromechanical and microstructural characteristics. Initially, they are annealed at 120 $^\circ\text{C}$ for 4 h in a vacuum oven. They are then cooled down at $10 \text{ }^\circ\text{C h}^{-1}$ to room temperature. This thermal process is done in a vacuum oven to prevent the specimens' surfaces oxidation (Hirata et al. 1985). An Agilent 5500 atomic force microscope (AFM) and a Wyko Optical Profiler are used to measure the surface roughness of the

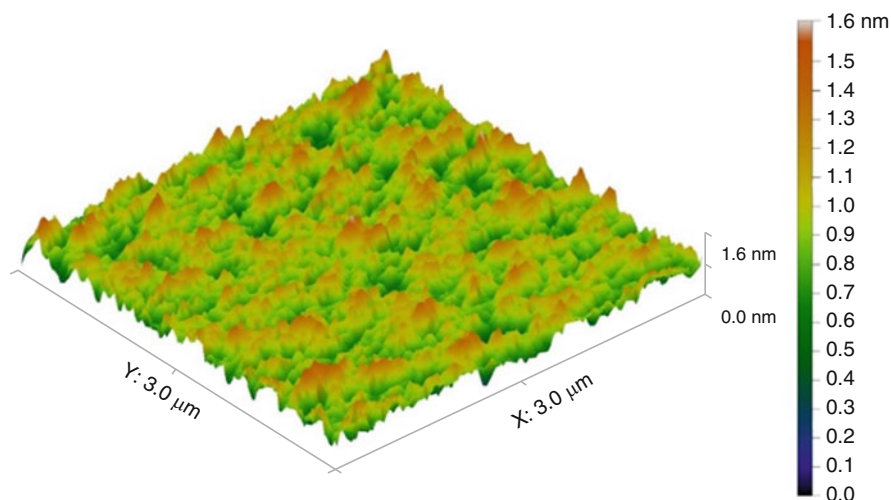


Fig. 1 Sample of AFM scanning of the PMMA sample surface (Reprinted from Malekmoitei et al. 2015)

samples, which markedly has effects on the nanoindentation experiment results (Kim et al. 2007; Nagy et al. 2013). The average surface roughness (R_a) of the PMMA and PC samples are measured to be 0.372 ± 0.013 and 0.305 ± 0.021 nm, respectively. Accordingly, since the samples' surfaces can be assumed as almost flat (Kim et al. 2007), no modification is needed for the obtained experimental results. Figure 1 represents a sample AFM scanning of the specimen surface. To indent the samples, they are mounted on aluminum stubs applying thermoresistant epoxy putty (Drummond™ Nu-Doh Epoxy Repair Compound Titanium Reinforced) for indentations at high temperature on hot-stage apparatus and super glue for indentations at ambient temperature.

Nanoindentation Technique

The nanoindentation technique has been utilized to measure the mechanical properties of the materials including the elastic modulus, E , and the hardness, H . The indentation load-hold-unload cycles have been performed using an MTS Nanoindenter® XP equipped with a three-sided pyramidal Berkovich diamond tip. The analysis of the applied load-indentation depth curves is mainly based on the original formalism developed by Oliver and Pharr (1992, 2004) as described in the following.

The hardness is described as the mean contact pressure under the indenter as follows:

$$H = \frac{P_{\max}}{A_c} \quad (1)$$

in which P_{\max} is the peak indentation load and A_c is the projected area of the tip-sample contact at the maximum load. According to Eq. 1, one needs an accurate measurement of the projected contact area under the load to evaluate the hardness from indentation load-displacement data. For a perfectly sharp Berkovich indenter, the contact area which is a function of the contact depth, h_c , can be calculated as

$$A_c = 24.56h_c^2 \quad (2)$$

For a practical indenter, which is not ideally sharp, the contact area function is required to be obtained by tip calibration with introducing additional terms to the aforementioned second-order relation as follows:

$$A_c = 24.56h_c^2 + C_1h_c^1 + C_2h_c^{1/2} + C_3h_c^{1/4} + \dots + C_8h_c^{1/128} \quad (3)$$

where C_1 through C_8 are constant coefficients which account for deviations from ideal geometry due to the blunting of the tip, and determined by using the indentation results on a standard fused silica sample and curve fitting performed on the Analyst[®] software. To obtain the exact contact area, an accurate determination of the depth over which the test material makes contact with the indenter, h_c , is required. The contact depth is generally different from the total penetration depth, and is estimated using

$$h_c = h - \varepsilon \frac{P}{S} \quad (4)$$

in which S is the measured elastic contact stiffness and ε is a constant which depends on the indenter geometry (for a Berkovich indenter $\varepsilon = 0.75$) (Oliver and Pharr 1992).

The elastic modulus of the sample, E , is calculated from Eq. 5 as follows:

$$\frac{1}{E_r} = \frac{1 - \nu^2}{E} + \frac{1 - \nu_i^2}{E_i} \quad (5)$$

In this equation, E_i and ν_i are the indenter elastic modulus and Poisson's ratio, respectively, ν is the sample Poisson's ratio, and E_r is the reduced elastic modulus. The reduced modulus which accounts for elastic deformation in both the indenter and the sample can be obtained from the following relation developed by Sneddon (1965)

$$E_r = \frac{S}{2\beta} \sqrt{\frac{\pi}{A_c}} \quad (6)$$

where β is a constant that depends on the indenter geometry and equals to 1.034 for the Berkovich tip.

Room-Temperature CSM Nanoindentation

The room temperature nanoindentation experiments are conducted by employing the continuous stiffness measurement (CSM) technique. The CSM method makes the continuous measurement of the mechanical properties of materials possible during the indentation loading segment from zero to the maximum indentation depth during a single test; this includes the measurement of the elastic contact stiffness at any point along the loading curve, and not just at the point of initial unload as in the basic mode (Pethica and Oliver 1988; Lucas et al. 1998; Li and Bhushan 2002; Hay et al. 2010). In this technique, a small sinusoidally varying load is superimposed on top of the primary loading signal that drives the motion of the indenter, and the resulting response of the system is analyzed by means of a frequency-specific amplifier. The displacement amplitude and the frequency of the superimposed oscillating force are set as 2 nm and 45 Hz, respectively, which are optimum values for the MTS nanoindentation[®] XP.

Utilizing the so-called CSM technique, the indentation load-hold-unload cycles are performed by keeping the loading rate divided by the load ratio (\dot{P}/P) constant during the loading segment over the course of a single indentation test. In a deep indentation test, where the indentation size effect (ISE) is negligible (i.e., the hardness value is almost constant), the loading path follows a Hertzian contact relation as $P = \alpha h^\beta$, where P represents the indentation load, h is the indentation depth, α is a material dependent parameter, and β is a curve fitting parameter close to 2 for the Berkovich tip (Johnson 1987; Zhang et al. 2005). Accordingly, for a pyramidal indenter, the indentation strain rate $\dot{\epsilon}_i$, which is defined as the descent rate of the indenter divided by its instantaneous depth, can be obtained by

$$\dot{\epsilon}_i = \frac{\dot{h}}{h} = \frac{1}{\beta} \frac{\dot{P}}{P} \quad (7)$$

Therefore, since the \dot{P}/P ratio is remained constant during the CSM test, at deep enough indentations, the indentation strain rate approaches a constant value equal to $\frac{1}{2} \frac{\dot{P}}{P}$ (Lucas and Oliver 1999; Voyiadjis and Malekmotiei 2016). Moreover, the effective shear strain rate induced by the indentation is then related to the indentation strain rate as follows:

$$\dot{\gamma} = \sqrt{3}C \dot{\epsilon}_i = \frac{\sqrt{3}C}{\beta} \frac{\dot{P}}{P} \quad (8)$$

where C is a constant equal to 0.09 (Poisl et al. 1995; Schuh and Nieh 2003; Schuh et al. 2004).

After samples preparation and calibration of the indenter tip function, the tests are carried out by typical loading-hold-unloading sequences as follows: prior to any measurement, the tip is held on the top of the sample surface until its drift rate is stabilized at a rate below 0.05 nms^{-1} . The tip then starts to travel downward

until it reaches the surface. Once the indenter tip comes into the contact with the sample surface, the loading segment begins with the constant value of \dot{P}/P until a predefined maximum depth of about 10 μm ; the load is then held at the maximum value for 10 s to account for the material creep behavior of the polymer surface; and finally, the unloading stage is carried out with the constant unloading rate until 10% of the maximum load is attained. Since one of the main purposes of this study is investigation of the strain rate dependency of the flow in PMMA and PC, a series of experiments are conducted at room temperature with the values of \dot{P}/P varying from 0.001–0.2 and 0.002–0.4 s^{-1} on annealed and as-cast PMMA and PC samples, respectively. A total of 25 indents are performed for each \dot{P}/P value with a minimum distance of 150 μm between neighboring indents to prevent from interaction.

High-Temperature Nanoindentation Procedure

The MTS Nanoindenter[®] XP is equipped with temperature control system for elevated temperature nanoindentation tests. The system includes a hot stage, a coolant apparatus to transfer the extra heat to the outside of the instrument, and a heat shield to keep the indenter transducer apart from the heat source. The load control experiments are performed by using the basic method which is employed for high-temperature indentations. To make sure that there is no indentation size effect, and making the obtained data comparable to those of the CSM tests as well, the experiments are carried out with a maximum load of 300 mN. Loading rates of 4, 10, 50, 100, and 300 mNs^{-1} for PMMA and 10, 100, and 300 mNs^{-1} for PC, which are constant during the tests, are applied for temperatures varying from room temperature to 100 and 140 $^{\circ}\text{C}$ (slightly lower than the samples glass transition temperature) for PMMA and PC, respectively.

While the loading rate \dot{P} is kept constant during the loading and unloading segments of the basic mode, evaluating the \dot{P}/P ratio shows that it changes with $1/h^{\beta}$ during the test since the load-depth ($P-h$) curve follows a Hertzian relation according to the statements in the preceding section. Therefore, the indentation strain rate $\dot{\epsilon}_i$ varies as $1/h^{\beta}$ during a test as well. Using Eqs. 7 and 8 to calculate the effective shear strain rate, an average value of \dot{P}/P over the deep part of the indentation (5 μm in this study) is considered in each test. Accordingly, the corresponding effective shear strain rates for the load rates given above are 0.0014, 0.0035, 0.0175, 0.035, and 0.105 s^{-1} for PMMA and 0.0035, 0.035, and 0.105 s^{-1} for PC, respectively.

Applying a thermoresistant epoxy putty (Drummond[™] Nu-Doh Epoxy Repair Compound Titanium Reinforced), the samples are mounted on the hot stage. Since polymer samples and the adhering thermoresistant have low thermal conductivity, the temperature of the sample surface can be considerably different from the set temperature. For precise evaluation, at the end of the tests, an Omega[®] SA1-K-SRTC thermocouple is attached to a sample surface to measure its temperature. The temperature is recorded by means of an Omega[®] HH74K handheld monitoring

device coupled with the thermocouple. The discrepancy between the measured and set temperatures is significant for high temperatures. The number of indentation points for each set temperature and loading rate is reduced to 12 for the elevated temperature tests to minimize the risk of the contamination of the tip by the softened polymer at high temperatures; the minimum distance of the adjacent indents is also increased to 200 μm . Prior to performing the loading cycle, each sample is heated to the set temperature inside the indenter; and then it is left to equilibrate for 2–3 h. On the onset of the test cycle, the tip is held at the distance of about 1 μm from the sample surface for about 10 min to adjust allowable thermal drift. This delayed contact is believed to help the tip to reach a thermal equilibrium with the sample.

Theory of Homogeneous Flow for Glassy Polymers

The inhomogeneous flow in amorphous solids is mediated by the formation of shear bands, whereas their homogeneous flow is triggered by the nucleation of STZs (Spaepen 1977). There are some important differences between inhomogeneous and homogenous flows of glassy solids. While the inhomogeneous flow is strain rate independent in a way that the flow stress does not change significantly with the strain rate, in a homogenous flow, the higher strain rate applied to the sample results in a considerable higher flow stress (Schuh and Nieh 2003; Schuh et al. 2007; Yang et al. 2007). Another significant distinction is the generation of multiple pop-ins in the load-displacement curves during the inhomogeneous flow which appears only in the nanoindentation experiments (Golovin et al. 2001; Zhang et al. 2005; Yang et al. 2007).

It has been shown that the flow behavior of metallic glasses is temperature dependent in a sense that there exists a transition from inhomogeneous to homogeneous flow at a certain temperature. The temperature at which the transition happens is also strain rate sensitive: the higher the applied strain rate, the higher the transition temperature (Yang et al. 2007). To investigate the flow nature in PMMA and PC as glassy polymers, the nanoindentation tests are performed on both as-cast and annealed samples. Figures 2 and 3 represent the variation of the hardness with temperature for as-cast and annealed PMMA and PC samples at different loading rates, respectively. Considering the direct relation between the hardness H and the flow stress σ_y as $H = \kappa \sigma_y$ with κ is Tabor's factor, these figures show that the hardness (or flow stress) is greatly strain rate sensitive in a way that a higher hardness is obtained at a higher strain rate (loading rate) at a given temperature. This observed rate-dependent softening, which is attributed to the thermally activated nature of the flow, indicates that the flow of PMMA and PC at tested temperatures, which are below their glass transition temperatures T_g , is homogeneous. In addition, Figs. 4 and 5 show the load-displacement (P - h) curves of the samples at the loading rates of 10 and 300 mNs^{-1} and different temperatures. All the curves are represented with the origin offset of 2 μm except the first one at room temperature. Since the curves are smooth with no pop-in events, Figs. 4 and 5 further confirm the

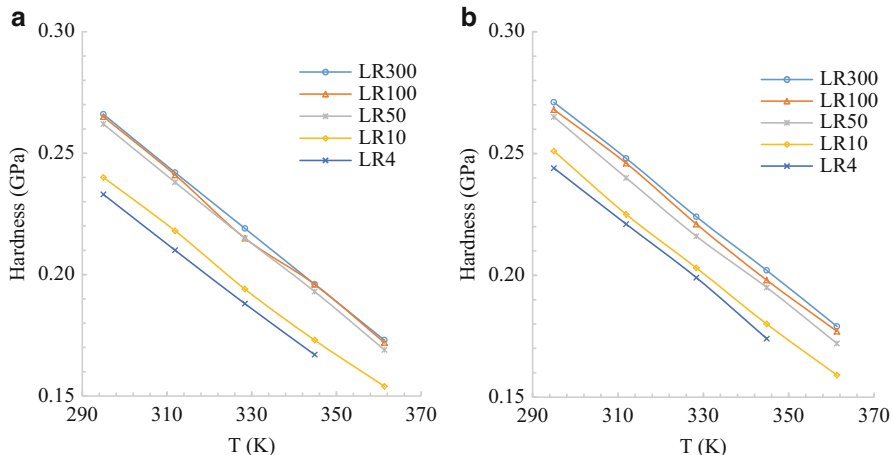


Fig. 2 Variation of the PMMA hardness with temperature at different loading rates for (a) as-cast and (b) annealed samples. The data points at the load rate of 4 mNs^{-1} and 361 K are not shown due to their high standard deviation values (Reprinted from Malekmoitei et al. 2015)

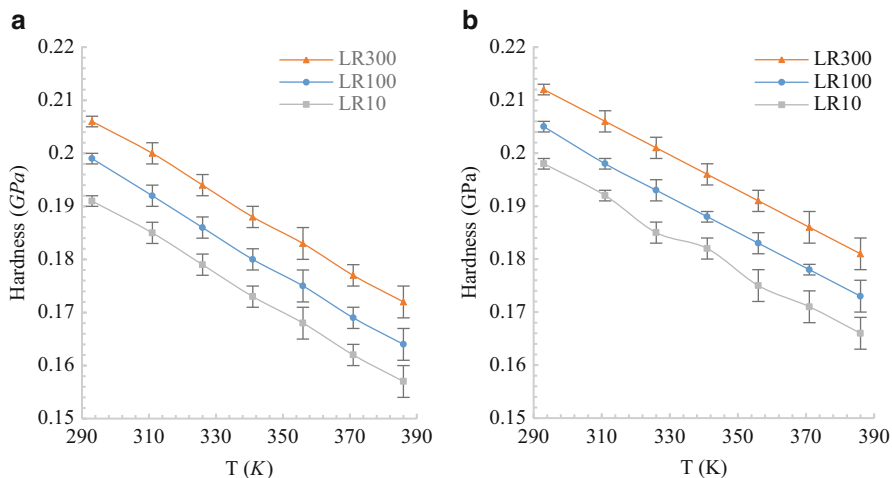
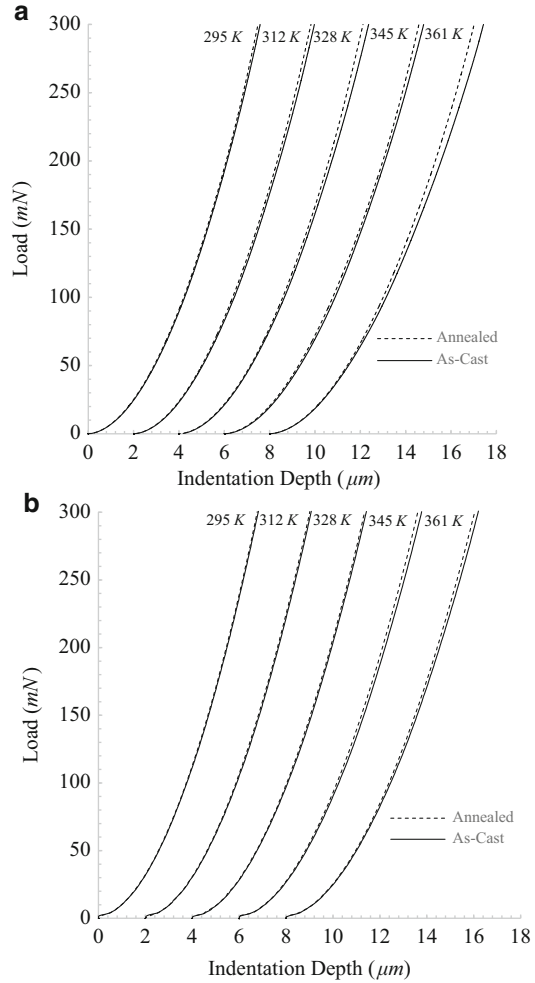


Fig. 3 Variation of the PC hardness with temperature at different loading rates for (a) as-cast and (b) annealed samples (Reprinted from Samadi-Dooki et al. 2016)

homogenous nature of the flow in PMMA and PC at tested temperatures and strain rates.

The flow in glassy solids is mediated by the irreversible local disturbances which form rearranged atomic (in MGs) or molecular (in PGs) clusters. These cooperative rearrangements result in isolated unit increments of shear, and are known as shear transformation zones (STZs). While in the homogeneous flow regime each volume element has contribution to the total plastic strain, the strain is localized in distinct

Fig. 4 Load-indentation depth curves for as-cast and annealed PMMA samples at different temperatures and the loading rates of (a) 10 mNs^{-1} , and (b) 300 mNs^{-1} (Reprinted from Malekmoitei et al. 2015)

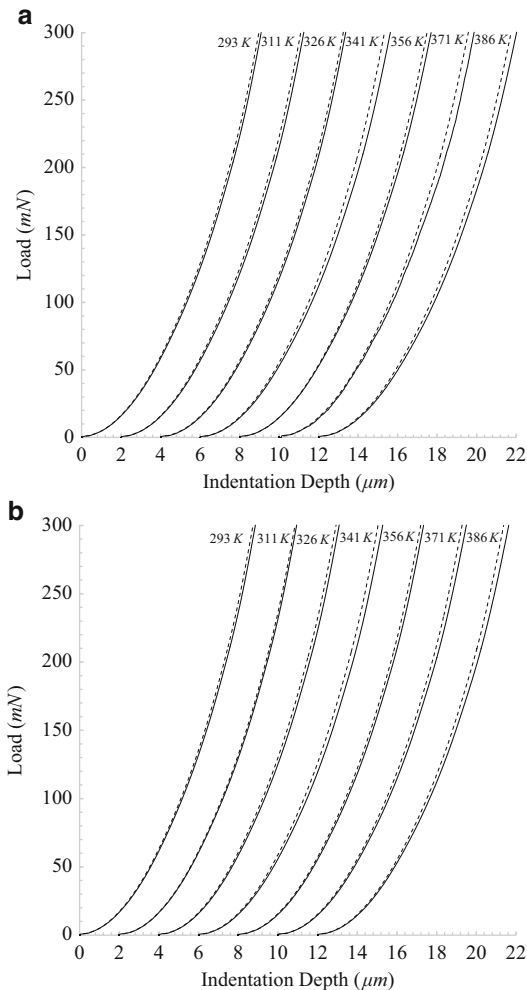


shear bands in the inhomogeneous flow regime (Spaepen 1977). The thermally activated homogeneous flow in glassy polymers can be described based on the flow mechanism developed by Spaepen (1977) and Argon (1979). For the STZs-mediated homogeneous flow in glassy solids, the kinetics relation for the shear strain rate $\dot{\gamma}$ due to the applied shear stress τ is well expressed by an Arrhenius relation as follows (Spaepen 1977; Argon 1979):

$$\dot{\gamma} = \dot{\gamma}_0 \exp\left(-\frac{\Delta F_0}{k_B T}\right) \sinh\left(\frac{\gamma^T \Omega \tau}{2k_B T}\right) \quad (9)$$

where $\dot{\gamma}_0$ is the pre-exponential factor proportional to the attempt frequency, k_B is the Boltzmann constant, T is the absolute temperature, and ΔF_0 is the nucleation

Fig. 5 The load-indentation depth curves for as-cast (*solid line*) and annealed (*dashed line*) PC samples at different temperatures and loading rates of (a) 10 mNs^{-1} and (b) 300 mNs^{-1} (Reprinted from Samadi-Dooki et al. 2016)



energy of an STZ with the shear strain γ^T occurring in a region of volume Ω . The factor 2 in the denominator of the argument of the hyperbolic function is due to the reverse transformation probability (Spaepen 1977). To evaluate the nucleation energy of an STZ, this locally transformed region has been treated as an embedded volume with nonelastic strain the micromechanical field of which can be obtained by using the Eshelby inclusion model (Eshelby 1957). For the RVE shown in Fig. 6, the free energy of the nucleation of a single STZ is given as follows:

$$\Delta F_0 = [\Xi(\nu) + \Psi(\nu)\beta^2] \mu (\gamma^T)^2 \Omega \quad (10)$$

in which μ is the shear modulus, $\Xi(\nu)$ and $\Psi(\nu)$ are functions of the Poisson's ratio ν , and pertain to the shear and dilatational components of the transformation strain

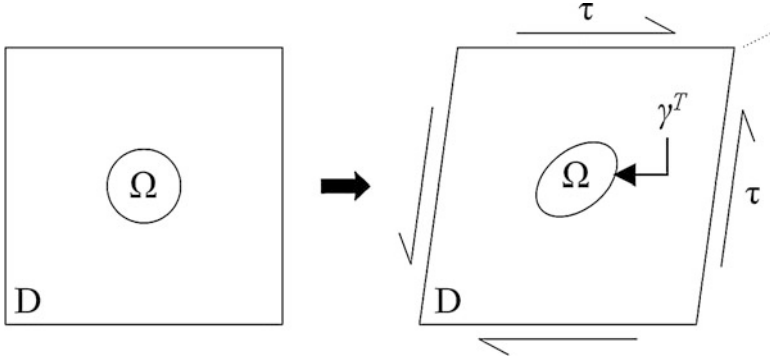


Fig. 6 Representative volume element (RVE) of polymer matrix containing a single shear transformation zone (STZ) (Reprinted from Malekmoitei et al. 2015)

tensor, respectively. The coefficient β is the dilatancy parameter and can be obtained from the pressure sensitivity of the flow. While $\Psi(\nu) = \frac{2(1+\nu)}{9(1-\nu)}$ is independent of the aspect ratio of the ellipsoid, $\mathcal{E}(\nu)$ is shape dependent (Mura 1987). The STZ's characteristics in Eq. 9, such as shear strain, size, and activation energy, have been experimentally obtained for different metallic glasses (see Chap. 7 of Argon (2013) for details), but their numerical evaluation for PGs have been limited to some theoretical modeling and simulations (Mott et al. 1993; Argon 2013), and a few experimental studies (Argon and Bessonov 1977; Ho et al. 2003).

Since $\gamma^T \Omega \tau \gg 2k_B T$ for conventional PGs at temperatures below their glass transition, Eq. 9 can be rearranged as:

$$\ln \dot{\gamma} = \frac{\gamma^T \Omega}{2k_B T} \tau + C_1 \quad (11)$$

where $C_1 = \ln \frac{\dot{\gamma}_0}{2} - \frac{\Delta F_0}{k_B T}$ represents a temperature-dependent parameter. In Eq. 11, the material constant $\gamma^T \Omega$, which is proportional to the important characteristic parameter of glassy polymers known as shear activation volume V^* (Ward 1971), can be determined from the derivative of the natural logarithm of the strain rate ($\ln \dot{\gamma}$) with respect to the shear flow stress (τ). Accordingly, considering the effect of hydrostatic pressure on the shear yield stress of polymers, the modified shear activation volume is obtained from Eq. 12 as follows (Ward 1971; Ho et al. 2003):

$$V^* = (1 - \alpha\beta) \gamma^T \Omega = 2k_B T (1 - \alpha\beta) \frac{\partial \ln \dot{\gamma}}{\partial \tau} \quad (12)$$

where β is the yield stress sensitivity to the pressure as defined before, and α is the loading condition constant which is between 0.6 and 0.7 for different compressive loading conditions (Ward 1971; Tervoort 1996). In fact, one can conclude from Eq. 12 that the shear activation volume is a modified STZ volume with considering

the dilatation effect. Furthermore, since γ^T is assumed to be a universal constant for all glassy polymers and equals about 0.02 (Mott et al. 1993; Ho et al. 2003), the size of the single shear transformation zone Ω can be obtained. Rearranging Eq. 9 for a constant strain rate, the shear flow stress can be expressed as a linear function of temperature as follows:

$$\tau = \Theta C_2 + \frac{2\Delta F_0}{\gamma^T \Omega} \quad (13)$$

where $C_2 = \frac{2k_B}{\gamma^T \Omega} \left(\ln \dot{\gamma} - \ln \frac{\dot{\gamma}_0}{2} \right)$ is representing a strain rate-dependent constant. As a result, the activation energy of an individual STZ, ΔF_0 , can be calculated from the linear interpolation of the variation of the flow shear stress with temperature by incorporating the obtained $\gamma^T \Omega$ values. Knowing all the STZ's parameters in Eq. 10, one can obtain the numerical value of $\mathcal{E}(\nu)$ and, consequently, an approximation of the STZ's shape in glassy polymers.

Results and Discussion

Calibrating the Nanoindentation Results

The instrumented-indentation testing (IIT) can be employed for the purpose of probing the mechanical properties including hardness and elastic modulus of very small volumes of materials including polymers, which both of them depend on the applied load on the sample surface and the contact area. In the nanoindentation experiments, the load is recorded with an nN scale precision while the contact area is calculated as a function of the tip geometry and contact depth. Since the tip in a practical indenter is not ideally sharp, the contact area function is required to be obtained by calibrating the tip with introducing a polynomial approximation (Oliver and Pharr 1992; Voyiadjis and Zhang 2015). Using the results of the indentation on standard fused silica sample, the tip areal function is calibrated to obtain reliable results.

It is also worth noting that the material pile-up around the penetrating tip can significantly alter the contact area and affect the measured mechanical properties of the material especially at shallow depths of the indentation. Using the optical profiler, the pile-up values are precisely measured for the indentation on PMMA and PC samples. As can be seen in Fig. 7, which shows a sample of pile-up measurement on PC specimen, the pile-up is highly unsymmetrical with the maximum values around the pyramidal tip faces. The maximum pile-up is measured to be about 400 nm, which in comparison with the maximum indentation depth of 10 μm is very small. For that reason, the pile-up effect is neglected in the nanoindentation measurements.

Another important observation during the nanoindentation experiments is the increment of the hardness values at shallow indentation depths which is known as

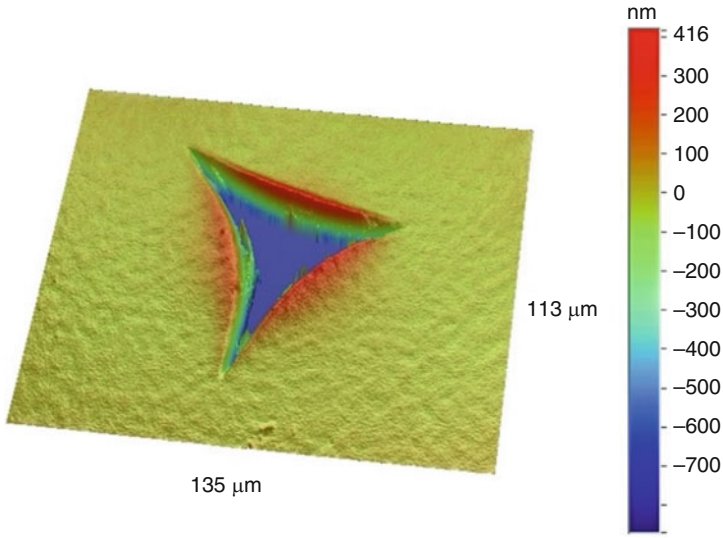


Fig. 7 A sample of the indentation pile-up measurement on PC specimen using Wyko Optical Profiler (Reprinted from Samadi-Dooki et al. 2016)

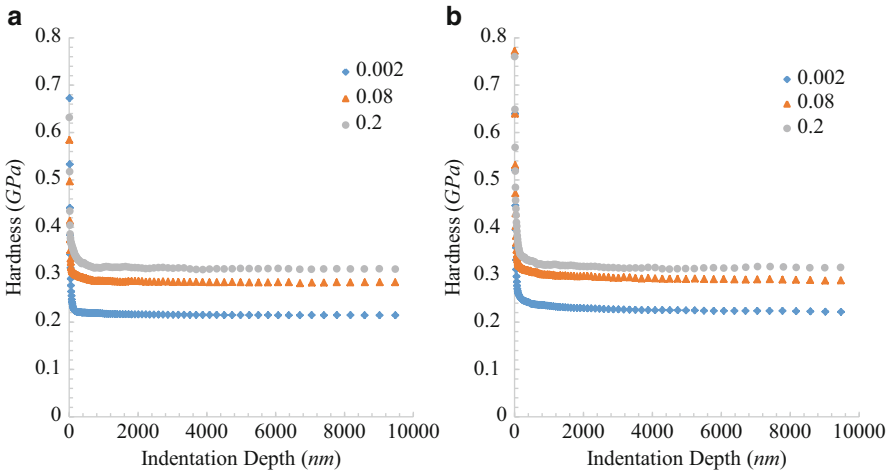


Fig. 8 Variation of the hardness versus the indentation depth for three different \dot{P}/P values for (a) as-cast and (b) annealed PMMA samples. The legend numbers represent the \dot{P}/P values (Reprinted from Malekmotiei et al. 2015)

indentation size effect (ISE) (Briscoe et al. 1998; Voyiadjis and Zhang 2015). To prevent the ISE in the current study, the CSM indentation results are first evaluated to find the indentation depth beyond which the obtained hardness reaches the stable value. The variation of hardness versus the indentation depth for both annealed and

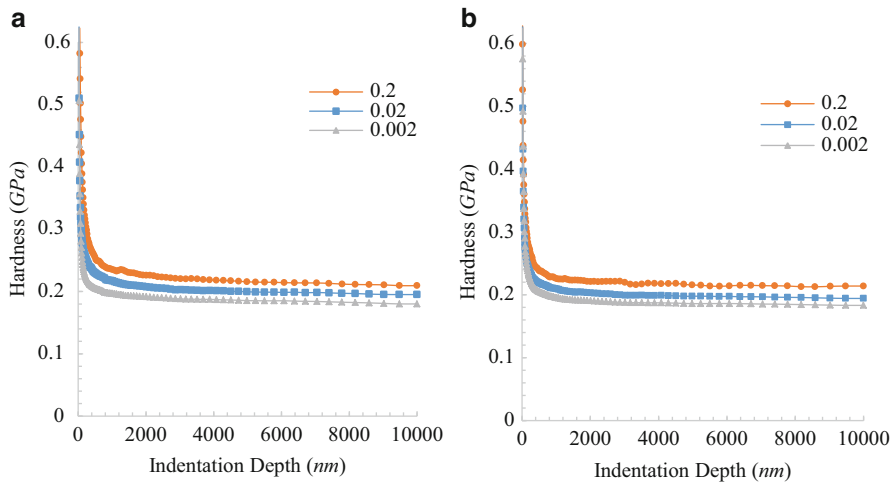


Fig. 9 Variation of the hardness versus the indentation depth for three different \dot{P}/P values for (a) as-cast and (b) annealed PC samples. The legend numbers represent the \dot{P}/P values (Reprinted from Samadi-Dooki et al. 2016)

as-cast PMMA and PC samples are presented in Figs. 8 and 9, respectively. As these figures show, the profound ISE is completely eliminated for the indentation depth beyond 2 and 4 μm for PMMA and PC samples, respectively. Accordingly, the hardness values are averaged over the indentation depth beyond 5 μm where the indentation size effect is absent. The elevated temperature tests are performed by using the basic mode in which the hardness value is only reported at the maximum indentation depth. While the maximum indentation depth is the input for the CSM technique, the maximum load is the input in the basic mode. As mentioned before, to avoid the indentation size effect, and also make the obtained data comparable to those of the CSM tests, the high temperature experiments are carried out with a maximum load of 300 mN, which corresponds to the maximum indentation depth of about 10 μm . Furthermore, since the ISE reduces at elevated temperatures (Voyiadjis et al. 2011), it is assured that the obtained hardness results are within the stable region.

Shear Activation Volume

In addition to the hardness, the elastic modulus is another mechanical property of the material which is continuously recorded during the loading segment of the CSM nanoindentation as a function of displacement. Figures 10 and 11 represent the variation of the material elastic modulus versus the indentation depth for some selected \dot{P}/P values for PMMA and PC samples, respectively. In comparison to the hardness, the elastic modulus is almost constant for the indentation depth beyond about 100 and 150 nm for both as-cast and annealed PMMA and PC samples,

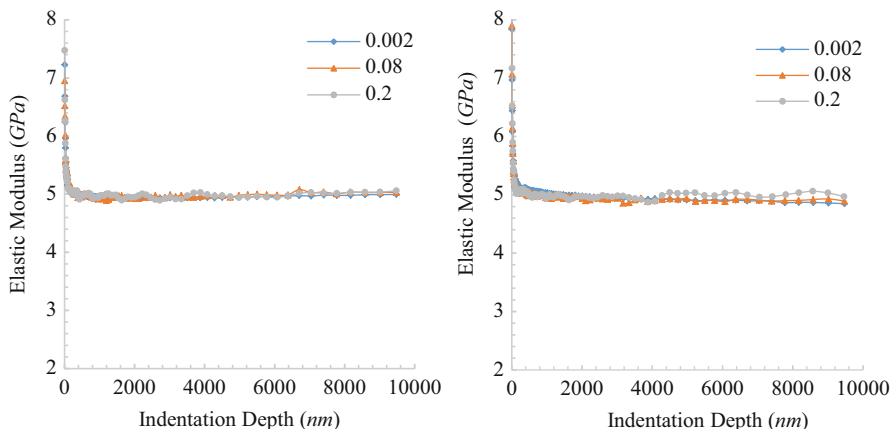


Fig. 10 Variation of the elastic modulus versus the indentation depth for three different \dot{P}/P values for (a) as-cast and (b) annealed PMMA samples. The legend numbers represent the \dot{P}/P values (Reprinted from Malekmotiei et al. 2015)

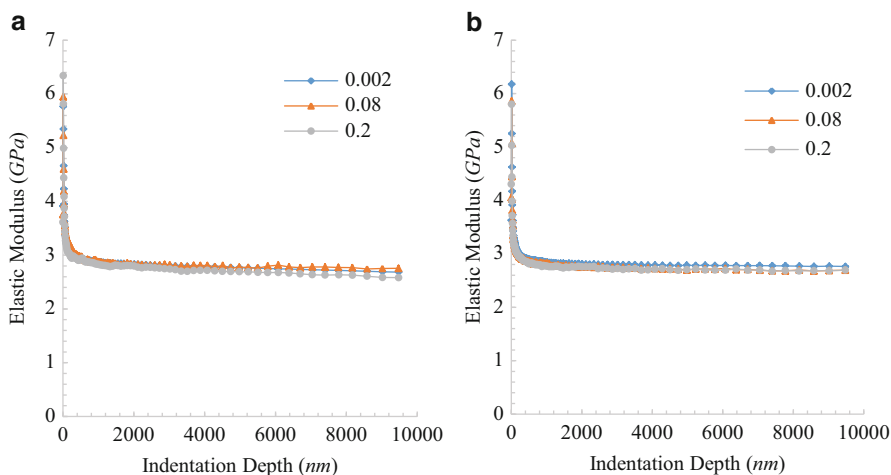


Fig. 11 Variation of the elastic modulus versus the indentation depth for three different \dot{P}/P values for (a) as-cast and (b) annealed PC samples. The legend numbers represent the \dot{P}/P values (Reprinted from Samadi-Dooki et al. 2016)

respectively. More importantly, while the hardness values significantly change with the strain rate (see Figs. 8 and 9), the elastic modulus does not show any significant strain rate sensitivity.

Using the Tabor's factor, the hardness values H obtained from the nanoindentation experiments can be converted to the yield stress σ_y of the material as follows:

$$H = \kappa \sigma_y \quad (14)$$

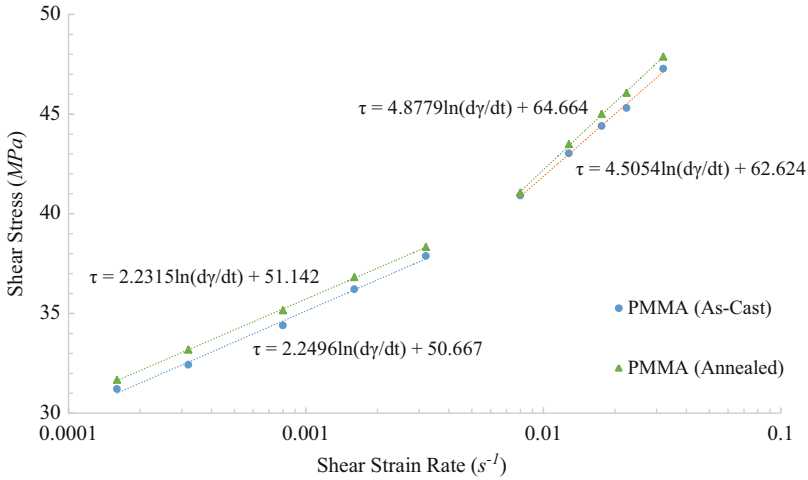


Fig. 12 Variation of the shear flow stress with the shear strain rate for both as-cast and annealed PMMA samples (Reprinted from Malekmotiei et al. 2015)

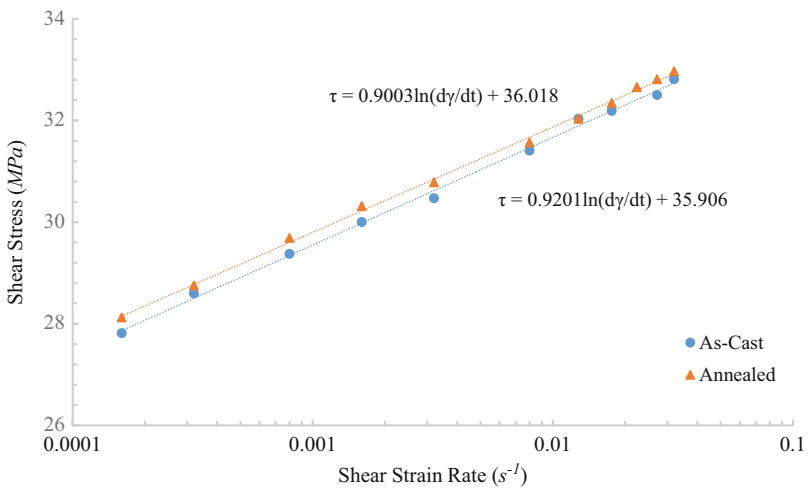


Fig. 13 Variation of the shear flow stress with the shear strain rate for both as-cast and annealed PC samples (Reprinted from Samadi-Dooki et al. 2016)

in which κ is Tabor's factor which is approximately 3.3 for amorphous polymers at high indentation strains (Prasad et al. 2009). Additionally, since the shear flow stress is about half of the yield stress in plane stress condition for monotonic loading (Yang et al. 2007), the ratio of the hardness to the shear flow stress is about 6.6. Figures 12 and 13 illustrate the shear flow stress-shear strain rate data

points for the as-cast and annealed PMMA and PC samples measured by CSM method at room temperature, respectively. The obtained results for PC are in good agreement with the ones obtained by Bauwens-Crowet et al. (1969) at 21.5 °C. As it is expected, the annealed samples have slightly bigger shear flow stresses compared to the as-cast ones at the same shear strain rate (Jancar et al. 2013). The outstanding feature of Fig. 12 is the existence of a significant transition at a certain value of the strain rate beyond which the strain rate sensitivity of the shear flow stress increases. This phenomenon is believed to be a result of strain rate shift of the β -relaxation process in the storage modulus of the PMMA, which is related to the restriction of the ester side group rotations at high strain rates, besides the intermolecular and local back bone motion restrictions (Calleja et al. 1994; Mulliken and Boyce 2006; Argon 2013). As obviously shown in Fig. 12, the transition shear strain rate is approximately 0.005 s^{-1} for both as-cast and annealed PMMA samples. Following the descriptions presented in Mulliken and Boyce (2006), the flow stress regimes below and above the transition shear strain rate might be referred to as α and β regimes, respectively. However, it is noteworthy to mention that since the room temperature β -transition strain rate of PC has been previously detected to be about 10^2 s^{-1} (Mulliken and Boyce 2006), which is beyond the strain rates that can be applied in nanoindentation experiments, no considerable jump is observed in Fig. 13 for the range of strain rates in this study.

Based on Eqs. 11 and 12, the shear activation volume V^* for an amorphous polymer can be obtained by linear interpolation of the $\tau - \ln \dot{\gamma}$ curve. In these figures, the slopes of the semilogarithmic stress-strain rate plots are almost the same for samples with different thermal history, which suggest that the shear activation volume and, therefore, the size of a single STZ are almost independent of the thermal history of the samples. Incorporating Eq. 11 and the data represented on Figs. 12 and 13, the factor $\gamma^T \Omega$ for as-cast and annealed samples is obtained about 3.66 and 3.69 nm^3 for PMMA, and 8.94 and 9.14 nm^3 for PC, respectively. Accordingly, by assuming $\alpha = 0.65$, $\beta = 0.204$ for PMMA (Ward 1971) and 0.27 for PC (Rittel and Dorogoy 2008), V^* is found to be 3.17 and 3.20 nm^3 in α regime for PMMA, and 7.37 and 7.54 nm^3 for PC, for as-cast and annealed samples, respectively. The obtained values of shear activation volume are in consonance with the molecular dynamics simulation results (Argon 2013). As results show, the shear activation volume for samples with different thermal histories is almost the same, and this small discrepancy might be due to the short time of annealing in this study (4 h). Since it has been shown that the flow stress of glassy polymers increases logarithmically with the annealing time at temperatures below their glass transitions (Hutchinson et al. 1999), a more profound difference might be expected between the shear flow stress results of the as-cast and annealed samples in Figs. 12 and 13 for longer annealing time. However, the increased difference may or may not result in a considerable difference in the shear activation volume since it depends on the slope of the $\tau - \ln \dot{\gamma}$ plots and not the shear flow stress solely.

STZ's Activation Energy

In light of Eq. 13, linear interpolation of the flow stress as a function of the temperature $\tau-T$ can be used to obtain the STZ's activation energy. Figures 14 and 15 represent the variance of the shear flow stress of PMMA and PC with temperature, respectively, for different strain rates which is well interpolated with linear functions at each loading rate. Assuming the parameter $\gamma^T \Omega$ does not vary with temperature, the STZ's activation energy for both as-cast and annealed samples can be calculated at different shear strain rates as shown in Fig. 16 for PMMA. One of the most important features of this figure is the existence of jump in the activation energy at the strain rate range of $0.0035-0.0175 \text{ s}^{-1}$ which is consistent with the β -transition strain rate obtained from room-temperature CSM nanoindentation. Therefore, this jump might be referred to as the β -transition activation energy. Although the discrepancy in the activation energy for the annealed and as-cast samples is small for strain rates above the β -transition strain rate, the difference is profound for strain rates below this transition. Since the annealed PMMA sample is expected to have more ordered chains in comparison to the as-cast one, the slip and rotation of these chains are more restricted in this sample; consequently, the STZ's activation energy increases. In contrast, beyond the β -transition strain rate, the rapid loading does not allow the chains to rotate or slip smoothly which puts the annealed and as-cast samples in the same deformation condition, and as a result, the activation energy of STZs for high strain rates is approximately identical for samples with different thermal histories.

Another important feature of Fig. 16 is that the β -transition activation energy is much bigger for the as-cast PMMA than the annealed one (almost three times).

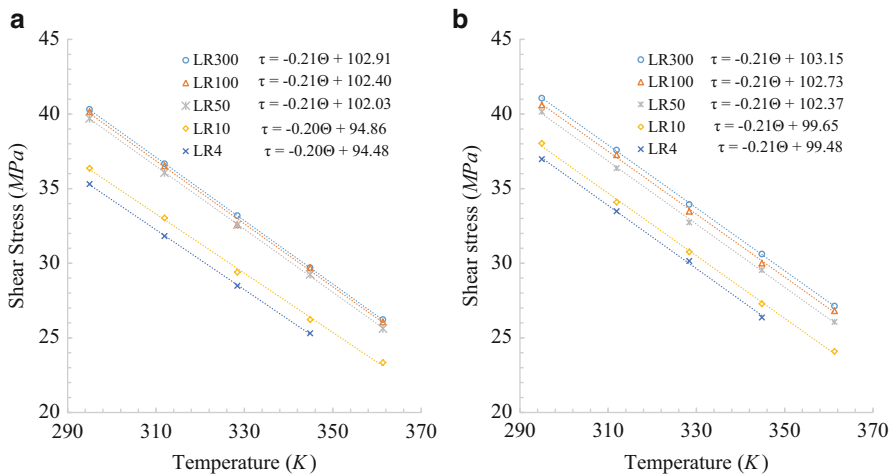


Fig. 14 Variation of the shear flow stress with temperature at different loading rates for (a) as-cast and (b) annealed PMMA samples (Reprinted from Malekmoetie et al. 2015)

Furthermore, the β -transition energy is about one order of magnitude smaller than the thermal activation energy of an STZ for PMMA, which is in agreement with the findings of Barral et al. (1994) who found almost the same ratio for a system containing a diglycidyl ether of bisphenol A (DGEBA) and 1,3-bisaminomethylcyclohexane (1,3-BAC). In comparison, the β -transition activation energy has been obtained to be almost equal to the STZ's activation energy in the metallic glasses (Yu et al. 2010).

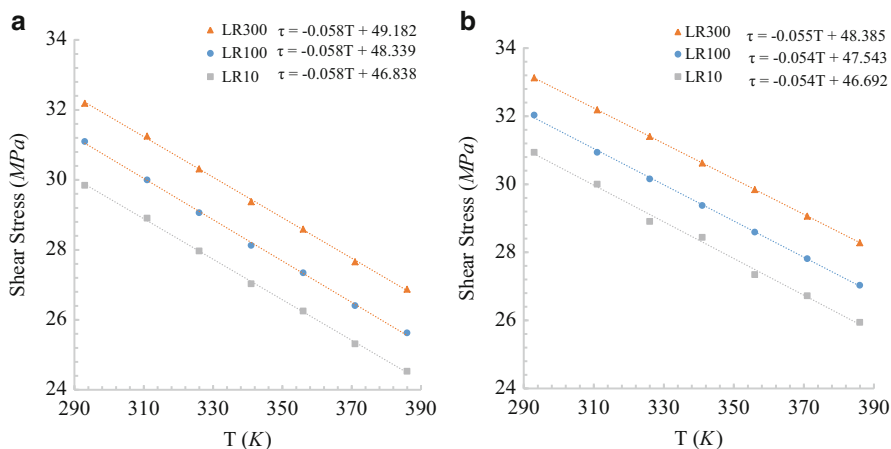


Fig. 15 Variation of the shear flow stress with temperature at different loading rates for (a) as-cast and (b) annealed PC samples (Reprinted from Samadi-Dooki et al. 2016)

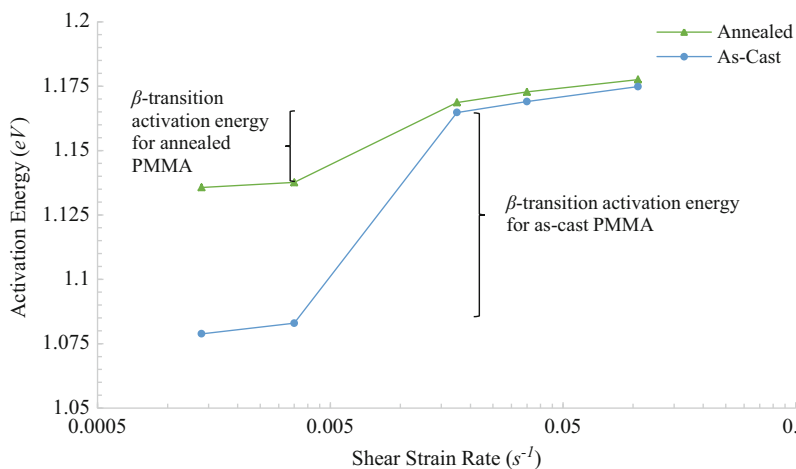


Fig. 16 STZ's activation energy for both as-cast and annealed PMMA samples at different shear strain rates (Reprinted from Malekmtouei et al. 2015)

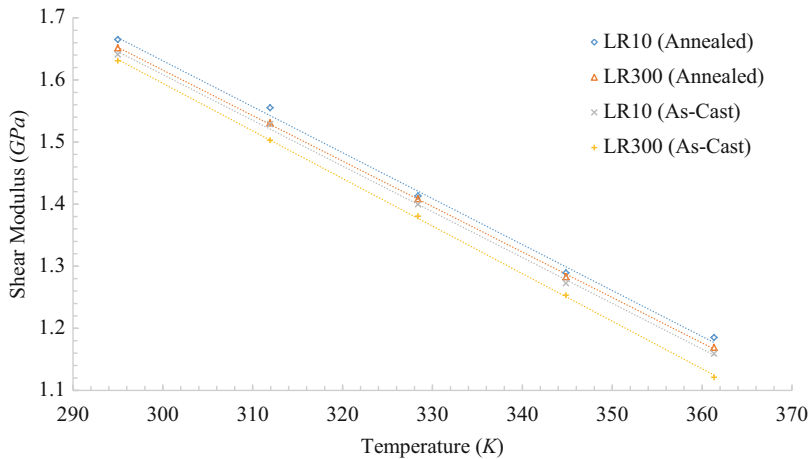


Fig. 17 Variation of the shear modulus with temperature for as-cast and annealed PMMA samples at two different loading rates (Reprinted from Malekmotiei et al. 2015)

As defined in Eq. 10, the Helmholtz free energy also depends on the shear modulus μ of the material which itself is temperature dependent. Since the shear modulus of the PMMA does not change considerably for temperatures below about 100 K (Gall and McCrum 1961), Eq. 13 is still valid, and ΔF_0 might be referred to as activation energy at 0–100 K. Using the elastic modulus data for different temperatures, obtained by basic method nanoindentation, and the relation between the elastic modulus and shear modulus as: $\mu = \frac{E}{2(1+\nu)}$, the variation of the shear modulus with temperature for both as-cast and annealed PMMA samples at two different strain rates are represented in Fig. 17. Assuming the variation to remain linear for temperatures down to 100 K, and constant for temperatures below 100 K which is a valid assumption based on the experiments of Gall and McCrum (1961), the variation of the STZ's activation energy with temperature can be obtained as shown in Fig. 18. It is also noticeable that the STZ's activation energy is around 0.6 eV at room temperature which is about one-third of that for metallic glasses (Yu et al. 2010).

Doing the same calculations on the obtained results for PC samples, the activation energy of a single STZ is presented in Table 1. As expected, the activation energy is slightly bigger for the annealed samples, and is almost strain rate insensitive which is in agreement with continuum mechanics principles. Since the obtained activation energies are calculated using the extrapolation of the $\tau-T$ values to the 0 K, they should be referred to as the zero Kelvin STZ's activation energies, and are shown by ΔF_0 hereafter. It should be mentioned that since the room temperature β -transition strain rate of PC is beyond the strain rates in this study, the β -transition activation energy barrier cannot be captured for PC samples.

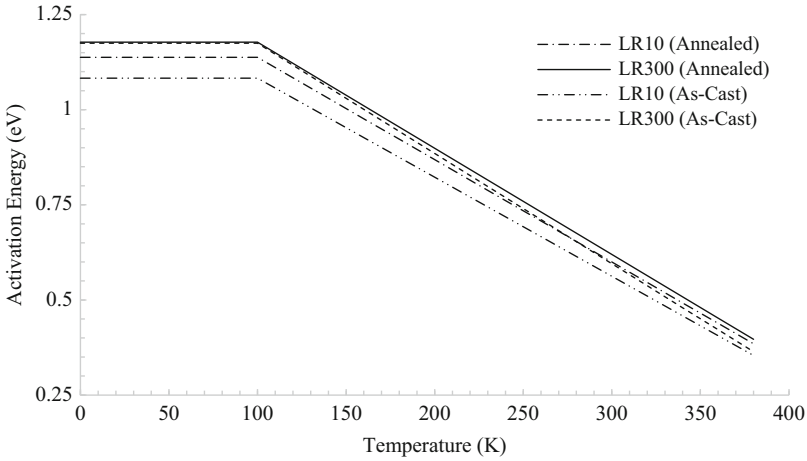


Fig. 18 Variation of the STZ’s activation energy with temperature for as-cast and annealed PMMA samples at two different loading rates (Reprinted from Malekmotiei et al. 2015)

Table 1 Characteristic properties of the STZs in PC for samples with different thermal histories (Reprinted from Samadi-Dooki et al. 2016)

Thermal history	$\gamma^T \Omega (nm^3)$	V^* (nm ³)	Activation energy (eV) at the shear strain rate of			γ^T	$\Omega (nm^3)$	Number of monomers
			0.0035 (s ⁻¹)	0.035 (s ⁻¹)	0.105 (s ⁻¹)			
As-cast	8.94	7.37	1.32	1.35	1.37	0.019	470	5,596
Annealed	9.14	7.54	1.33	1.36	1.38	0.0187	488	5,810

STZ’s Geometry

Incorporating Eqs. 10 and 13, the intercept of $\tau-T$ linear interpolation with τ axis is equal to $\frac{2\Delta F_0}{\gamma^T \Omega} = 2 [\Xi(v) + \Psi(v) \beta^2] \mu \gamma^T$. As it is mentioned before, $\Psi(v) = \frac{2(1+v)}{9(1-v)}$ does not vary with the STZ shape, and is about 0.5 and 0.48 for PMMA and PC, respectively (considering the Poisson’s ratio of 0.38 for PMMA and 0.37 for PC). However, $\Xi(v)$ is determined by the shape of the STZ which varies between $\frac{7-5v}{30(1-v)} = 0.27$ and 0.5 pertaining to the spherical and flat ellipsoidal shape of STZ, respectively (Mura 1987). Taking $\mu_0 = 3.1$ GPa from Fig. 17 and $\beta = 0.204$, the value of γ^T varies between 0.03 and 0.05 for PMMA, with the lower and upper bonds pertaining to the flat ellipsoidal and spherical transformation zones, respectively. Based on the molecular dynamics simulation (Mott et al. 1993), the shear strain (γ^T) of 0.05 is a large value for polymeric glasses; therefore, the shape of the transformation zone is more likely to be a flat ellipsoid for PMMA instead

of a sphere. Using the previously obtained factor $\gamma^T \Omega$ and assuming $\gamma^T = 0.03$, the volume of an individual plastic deformation unit Ω , which is almost the same for as-cast and annealed PMMA samples, is obtained about 123 nm^3 which is at least one order of magnitude bigger than that of the metallic glasses (Pan et al. 2008; Yu et al. 2010). Assuming the PMMA monomers as cylinders with the length of 1.55 \AA and radius of 2.85 \AA (Argon and Bessonov 1977), the single STZ is found to contain about 3000 monomers.

In a same way, using $\mu_0 = 2.4 \text{ GPa}$ (Argon 2013) and $\beta = 0.27$ (Rittel and Dorogoy 2008), γ^T is obtained equal to 0.035 for the spherical and about 0.02 for the flat ellipsoidal shapes of the STZ for both as-cast and annealed PC samples. Due to the aforementioned reason, it can be inferred that the STZs in PC are regions with the shapes close to flat ellipsoids rather than spheres. In addition, the volume of an STZ, Ω , is calculated to be about 470 and 488 nm^3 which has in average 5,600 and 5,800 idealized cylindrical shape monomers with 2.8 \AA length and 3.09 \AA radius (Argon and Bessonov 1977) for the as-cast and annealed PC samples, respectively. All the obtained characteristic properties of a single STZ in PC samples are presented in Table 1.

Concluding Remarks

In summary, nanoindentation experiments conducted on as-cast and annealed specimens are studied to explore the temperature and strain rate sensitivity of the flow in poly(methyl methacrylate) (PMMA) and polycarbonate (PC) as amorphous polymers. Showing that the flow is homogeneous in these polymers at temperatures below their glass transitions and incorporating a homogeneous flow theory, geometrical and micromechanical characteristics of the shear transformation zones (STZs), as main carriers of plasticity in amorphous polymers, have been investigated in the molecular level. Since the experimental studies on the STZs in glassy polymers is less addressed in the literature in comparison to the metallic glasses, the nanoindentation technique is employed as an accurate, repeatable, and nondestructive method to answer some of the current open questions in this area. The findings suggest that the STZs are flat ellipsoidal regions with the volume of about 123 and 480 nm^3 and the transformation shear strain of about 0.03 and 0.02 in PMMA and PC samples, respectively. In addition, the nucleation energy of the shear transportation zones for both samples as well as the β -transition energy barrier for PMMA samples has been obtained. The procedure used for obtaining the β -transition energy based on the nanoindentation technique is an innovative approach which can be used for other glassy solids. The experimentally evaluated parameters produce unequivocal values as inputs for further theoretical and numerical investigations of yielding and plasticity in polymeric glasses.

In light of the obtained results for PMMA and PC, the following remarks are made:

1. The STZ's nucleation energy in amorphous polymers is about 1 eV which is not considerably smaller than that of the metallic glasses. Since the nucleation energy of an STZ directly depends on the material shear modulus, this energy is expected to be much higher for MGs due to their considerably bigger shear modulus. However, the nucleation energy also depends on the STZ size. Therefore, the bigger size of the STZs in PGs compared to MGs compensates for their shear modulus discrepancy, and levels the activation energy in two materials.
2. The transformation shear strain γ^T is slightly bigger than what is believed to be the universal value for PGs. The transformation shear strain has been considered to be about 0.015 in all types of glassy polymers (Argon 2013); however, the current study suggests that this parameter is unique to a particular polymer, and is about 0.02 in PC and 0.03 in PMMA.
3. While this work suggests that the shear transformation zones are formed in the regions with the shape close to the flat ellipsoid in PMMA and PC, the STZ shape is assumed to be spherical in all types of glassy solids (Ho et al. 2003; Argon 2013; Li et al. 2013).
4. Since all experiments are performed at temperatures beyond 0.6 Tg, the observed homogeneous flow is in accordance with the amorphous flow theory at elevated temperatures (Argon 1979, 2013). With the current nanoindentation technology, it seems impossible to perform experiments at temperatures below 0.6 Tg for available PGs (−43 and −20 °C for PMMA and PC, respectively). Undoubtedly, experiments at temperatures below 0.6 Tg would result in a better understanding of the flow nature in glassy polymers at a wider temperature range.
5. The obtained results for the STZs shape are based on the acceptable values of transformation shear strain in polymers which is considerably smaller than that in metallic glasses. A precise evaluation requires the STZs direct observation, which is not possible and convenient since STZs are local transition events rather than being actual defects like dislocations, or their indirect observation via localized stress field monitoring, which to the best of the authors' knowledge has not been reported yet. The only indirect experimental measurement of the STZ's size scale is one by Liu et al. (2011) in which the 2.5 nm size of the viscoelastic heterogeneities observed by transmission electron microscopy (TEM) has been related to the size of the STZ in a metallic glass.

References

- L. Anand, M.E. Gurtin, *Int. J. Solids Struct.* **40**, 1465 (2003)
A. Argon, *Philos. Mag.* **28**, 839 (1973)
A. Argon, *Acta Metall.* **27**, 47 (1979)
A.S. Argon, *Mater. Sci. Technol.* (1993)
A. Argon, *Strengthening Mechanisms in Crystal Plasticity* (Oxford University Press on Demand, Oxford, 2008)
A.S. Argon, *The Physics of Deformation and Fracture of Polymers* (Cambridge University Press, Cambridge, 2013)
A.S. Argon, M. Bessonov, *Philos. Mag.* **35**, 917 (1977)

- E.M. Arruda, M.C. Boyce, R. Jayachandran, *Mech. Mater.* **19**, 193 (1995)
- L. Barral, J. Cano, A. López, P. Nogueira, C. Ramírez, *J. Therm. Anal. Calorim.* **41**, 1463 (1994)
- C. Bauwens-Crowet, J. Bauwens, G. Homes, *J. Polym. Sci. Part A-2: Polym. Phys.* **7**, 735 (1969)
- M.C. Boyce, D.M. Parks, A.S. Argon, *Mech. Mater.* **7**, 15 (1988)
- B. Briscoe, L. Fiori, E. Pelillo, *J. Phys. D. Appl. Phys.* **31**, 2395 (1998)
- R.D. Calleja, I. Devine, L. Gargallo, D. Radić, *Polymer* **35**, 151 (1994)
- K. Chen, K.S. Schweizer, *Macromolecules* **44**, 3988 (2011)
- J.D. Eshelby, The determination of the elastic field of an ellipsoidal inclusion, and related problems. *Proc. R. Soc. Lond., Ser. A* **241**, 376 (1957)
- M.L. Falk, *Science* **318**, 1880 (2007)
- M.L. Falk, J.S. Langer, L. Pechenik, Toward a Shear-Transformation-Zone Theory of Amorphous Plasticity. In: Yip S. (eds) *Handbook of Materials Modeling* (Springer, Dordrecht, 2005), pp 1281–1312
- W. Gall, N. McCrum, *J. Polym. Sci.* **50**, 489 (1961)
- Y.I. Golovin, V. Ivolgin, V. Khonik, K. Kitagawa, A. Tyurin, *Scr. Mater.* **45**, 947 (2001)
- J. Hay, P. Agee, E. Herbert, *Exp. Tech.* **34**, 86 (2010)
- T. Hirata, T. Kashiwagi, J.E. Brown, *Macromolecules* **18**, 1410 (1985)
- J. Ho, L. Govaert, M. Utz, *Macromolecules* **36**, 7398 (2003)
- J. Hutchinson, S. Smith, B. Horne, G. Gourlay, *Macromolecules* **32**, 5046 (1999)
- J. Jancar, R.S. Hoy, A.J. Lesser, E. Jancarova, J. Zidek, *Macromolecules* **46**, 9409 (2013)
- K.L. Johnson, *Contact Mechanics* (Cambridge University Press, Cambridge, 1987)
- J. Ju, D. Jang, A. Nwankpa, M. Atzmon, *J. Appl. Phys.* **109**, 053522 (2011)
- J.-Y. Kim, S.-K. Kang, J.-J. Lee, J.-i. Jang, Y.-H. Lee, D. Kwon, *Acta Mater.* **55**, 3555 (2007)
- X. Li, B. Bhushan, *Mater. Charact.* **48**, 11 (2002)
- L. Li, E. Homer, C. Schuh, *Acta Mater.* **61**, 3347 (2013)
- Y. Liu, D. Wang, K. Nakajima, W. Zhang, A. Hirata, T. Nishi, A. Inoue, M. Chen, *Phys. Rev. Lett.* **106**, 125504 (2011)
- B. Lucas, W. Oliver, *Metall. Mater. Trans. A* **30**, 601 (1999)
- B. Lucas, W. Oliver, J. Swindeman, The dynamics of frequency-specific, depth-sensing indentation testing, *MRS Online Proceedings Library Archive* 522 (1998)
- L. Malekmotiei, F. Farahmand, H.M. Shodja, A. Samadi-Dooki, *J. Biomech. Eng.* **135**, 041004 (2013)
- L. Malekmotiei, A. Samadi-Dooki, G.Z. Voyiadjis, *Macromolecules* **48**, 5348 (2015)
- P. Mott, A. Argon, U. Suter, *Philos. Mag. A* **67**, 931 (1993)
- A. Mulliken, M. Boyce, *Int. J. Solids Struct.* **43**, 1331 (2006)
- T. Mura, *Micromechanics of defects in solids* (Springer Science & Business Media, 2013)
- P. Nagy, I.I. Kükemezey, S. Kassavetis, P. Berke, M.-P. Delplancke-Ogletree, S. Logothetidis, *Nanosci. Nanotechnol. Lett.* **5**, 480 (2013)
- E. Oleinik, S. Rudnev, O. Salamatina, *Polym. Sci. Ser. A* **49**, 1302 (2007)
- W.C. Oliver, G.M. Pharr, *J. Mater. Res.* **7**, 1564 (1992)
- W.C. Oliver, G.M. Pharr, *J. Mater. Res.* **19**, 3 (2004)
- D. Pan, A. Inoue, T. Sakurai, M. Chen, *Proc. Natl. Acad. Sci.* **105**, 14769 (2008)
- D. Pan, Y. Yokoyama, T. Fujita, Y. Liu, S. Kohara, A. Inoue, M. Chen, *Appl. Phys. Lett.* **95**, 141909 (2009)
- S. Pauly, S. Gorantla, G. Wang, U. Kühn, J. Eckert, *Nat. Mater.* **9**, 473 (2010)
- J. Pethica, W. Oliver, Mechanical properties of nanometre volumes of material: use of the elastic response of small area indentations, in *MRS Proceedings* (Cambridge University Press, 1988)
- W. Poisl, W. Oliver, B. Fabes, *J. Mater. Res.* **10**, 2024 (1995)
- K.E. Prasad, V. Keryvin, U. Ramamurty, *J. Mater. Res.* **24**, 890 (2009)
- T. Ree, H. Eyring, *J. Appl. Phys.* **26**, 800 (1955)
- D. Rittel, A. Dorogoy, *J. Mech. Phys. Solids* **56**, 3191 (2008)
- R.E. Robertson, *J. Chem. Phys.* **44**, 3950 (1966)
- A. Samadi-Dooki, L. Malekmotiei, G.Z. Voyiadjis, *Polymer* **82**, 238 (2016)
- C.A. Schuh, T. Nieh, *Acta Mater.* **51**, 87 (2003)

- C.A. Schuh, A.C. Lund, T. Nieh, *Acta Mater.* **52**, 5879 (2004)
- C.A. Schuh, T.C. Hufnagel, U. Ramamurty, *Acta Mater.* **55**, 4067 (2007)
- H. Shodja, I. Rad, R. Soheilifard, *J. Mech. Phys. Solids* **51**, 945 (2003)
- I.N. Sneddon, *Int. J. Eng. Sci.* **3**, 47 (1965)
- F. Spaepen, *Acta Metall.* **25**, 407 (1977)
- G. Spathis, E. Kontou, *J. Appl. Polym. Sci.* **79**, 2534 (2001)
- G. Stoclet, R. Seguela, J. Lefebvre, S. Elkoun, C. Vanmansart, *Macromolecules* **43**, 1488 (2010)
- G. Tandon, G. Weng, *Polym. Compos.* **5**, 327 (1984)
- T. A. Tervoort, *Constitutive modelling of polymer glasses: finite, nonlinear viscoelastic behaviour of polycarbonate* (Technische Universiteit Eindhoven, (1996)
- G.Z. Voyiadjis, L. Malekmoie, *J. Polym. Sci. Part B: Polym. Phys.* **54**, 2179 (2016)
- G.Z. Voyiadjis, A. Samadi-Dooki, *J. Appl. Phys.* **119**, 225104 (2016)
- G.Z. Voyiadjis, C. Zhang, *Mater. Sci. Eng. A* **621**, 218 (2015)
- G.Z. Voyiadjis, D. Faghihi, C. Zhang, *J. Nanomech. Micromech.* **1**, 24 (2011)
- I.M. Ward, *J. Mater. Sci.* **6**, 1397 (1971)
- B. Yang, J. Wadsworth, T.-G. Nieh, *Appl. Phys. Lett.* **90**, 061911 (2007)
- H. Yu, W. Wang, H. Bai, Y. Wu, M. Chen, *Phys. Rev. B* **81**, 220201 (2010)
- G. Zhang, W. Wang, B. Zhang, J. Tan, C. Liu, *Scr. Mater.* **52**, 1147 (2005)



Properties of Material Interfaces: Dynamic Local Versus Nonlocal

11

Devendra Verma, Chandra Prakash, and Vikas Tomar

Contents

Introduction	362
Interfaces	363
Interface Mechanical Properties	363
Interface Elastic Constants	366
Interface Dynamic Properties	367
Fracture Properties of Interfaces	369
Cohesive Zone Model	369
Interface Thickness Effect	370
Conclusions	374
References	375

Abstract

Interfaces in the materials are known entities since last century described as early as in the interfacial excess energy formulations by Gibbs (Boßelmann et al. 2007). The interface effect (or surface effect) is also widely referred to as the interface stress (or surface stress) that consists of two parts, both arise from the distorted atomic structure near the interface (or surface): the first part is the interface (or surface) residual stress which is independent of the deformation of solids, and the second part is the interface (or surface) elasticity which contributes to the stress field related to the deformation. Plastic deformation, in particular, the initial yielding point (i.e., the yield surface), is sensitive to the local stress (or local strain) of a heterogeneous material, which includes both the local (surface/interface) residual stress and local stress–strain relationship. The plastic

D. Verma · C. Prakash · V. Tomar (✉)

School of Aeronautics and Astronautics, Purdue University, West Lafayette, IN, USA

e-mail: tomar@purdue.edu

© Springer Nature Switzerland AG 2019

G. Z. Voyiadjis (ed.), *Handbook of Nonlocal Continuum Mechanics for Materials and Structures*, https://doi.org/10.1007/978-3-319-58729-5_21

361

deformation at the interfaces also considers the tension and compression along the interface and stress mismatch because of the material property differences. In the nanomaterials, the surface and interface stresses become even more important owing to the nanoscale size of the particles and interface areas.

Keywords

Nanomechanical Raman Spectroscopy · Interface · GB · CZM

Introduction

Interfaces in the materials are known entities since last century described as early as in the interfacial excess energy formulations by Gibbs (Boßelmann et al. 2007). The interface effect (or surface effect) is also widely referred to as the interface stress (or surface stress) that consists of two parts, both arise from the distorted atomic structure near the interface (or surface): the first part is the interface (or surface) residual stress which is independent of the deformation of solids, and the second part is the interface (or surface) elasticity which contributes to the stress field related to the deformation. Plastic deformation, in particular, the initial yielding point (i.e., the yield surface), is sensitive to the local stress (or local strain) of a heterogeneous material, which includes both the local (surface/interface) residual stress and local stress–strain relationship. The plastic deformation at the interfaces also considers the tension and compression along the interface and stress mismatch because of the material property differences. In the nanomaterials, the surface and interface stresses become even more important owing to the nanoscale size of the particles and interface areas.

The naturally occurring materials have been of keen interest in the materials community with the aim of understanding and reproducing the exceptional strength and toughening mechanisms present in exoskeletons of shrimps and lobsters (Boßelmann et al. 2007; Raabe et al. 2006), crabs (Mayer 2011; Chen et al. 2008), nacre (Flores-Johnson et al. 2014), etc. All these materials share some common traits such as a strong hierarchical structure, layered structure, composition of material with both minerals and fibers, gradient in the thickness of layers, etc. These naturally occurring materials have been able to manipulate the characteristics mentioned above to customize the design of their exoskeletons. The one important parameter along with the material composition in these designs is the role of the interfaces in the multilayered structure. In our articles, we highlighted the difference between the mechanical properties of two similar species of shrimp *Pandalus platyceros* and *Rimicaris exoculata* with interlayer structure as shown in Fig. 1a that are found at sea level and at 2300 m depth in the sea as a function of habitat, wet versus dry (Verma and Tomar 2014a), as a function of temperature (Verma and Tomar 2014b), and as a function of mineral composition (Verma and Tomar 2015a, b).

One unique feature that determines the properties of materials is the interfacial interaction between organic and inorganic phases in the form of protein (e.g., chitin (CHI) or tropocollagen (TC))-mineral (e.g., calcite (CAL) or hydroxyapatite

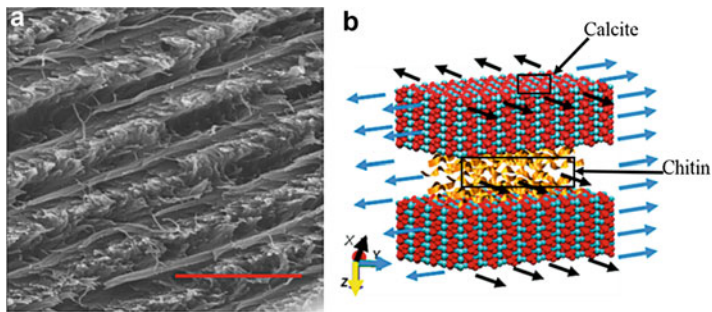


Fig. 1 (a) Interfacial structure in shrimp exoskeleton and (b) interfacial molecular structure made of Calcite and Chitin

(HAP)) interfaces. The size of the protein-mineral interfaces can be enormous as the mineral bits have nanoscale eliminations. These interfaces control biological reactions, and provide unique organic microenvironments that can enhance specific affinities, as well as self-assembly in the interface plane that can be used to orient and space molecules with precision. Interfaces also play a significant role in determining structural integrity and mechanical creep and strength properties of biomaterials. The length scale and complexity of microstructure of hybrid interfaces in biological materials make it difficult to study them and to understand the underlying mechanical principles, which are responsible for their extraordinary mechanical performance. One of the most important aspects of understanding the influence of interfaces on natural material properties is the knowledge of how stress transfer occurs across the organic-inorganic interfaces. Molecular modeling provides a way to study these phenomena at the length scale of the individual components. The effect of different kinds of interfaces was modeled by our research in the case of chitin layers. The interfaces with different thickness and different phases were compared to study the effectiveness of the stress transfer of the layers as shown in Fig. 1b. A comprehensive study on these is given in the articles by Tao et al. (Qu et al. 2015a, b).

In this article, a discussion is given on the interface mechanics. Nanomechanical Raman spectroscopy is used in this paper to measure the interface stresses. A comparison between the numerically and experimentally measured interface elastic constants is presented in addition to the stress-stress response of interface and its effect on material properties of the interface and its adjacent phases.

Interfaces

Interface Mechanical Properties

Interfaces in composite materials can be considered as a material phase confined between two separate grains or phases. Single interface samples of glass and epoxy were prepared with an epoxy interface sandwiched between two glass phases. The

samples were prepared using two-part industrial epoxy procured from Composite Polymer Design (South St. Paul, MN, USA).

The resin, CPD4505A, and hardener, CPD 4507B, were thoroughly mixed in recommended proportions of 100A: 28B by weight. The epoxy layer thickness was controlled by putting tabs of appropriate thickness in between the glass slides. The samples were cured at a prescribed temperature of 250 °F for 1 day.

The thickness of the interface in samples was measured with a microscope to make sure that it was in the error margin of $10 \pm 0.5 \mu\text{m}$. The sample surfaces as shown in Fig. 2a were polished to remove scratches that could interfere with the data measurement during experiments.

The Raman spectroscopy is based on “Raman Effect,” which provides a unique “fingerprint” of every individual substance as a characteristic for its identification. It is an inelastic process in which energy is exchanged between the incident photon and molecule. We have used the Raman spectroscopy to measure the stresses in the interface at different applied loads during indentation to compare the stress distribution. The Mechanical load was applied using the nanoindentation platform manufactured by Micro Materials Inc., UK, with load range from 0.1 to 500 mN, with the accuracy of better than 0.1 mN. The experimental setup to measure the Raman signal is shown in Fig. 2c.

The first step to measure the stress across the interface is to establish a calibration curve of Raman shift with the applied stress. A uniaxial load was applied on a block

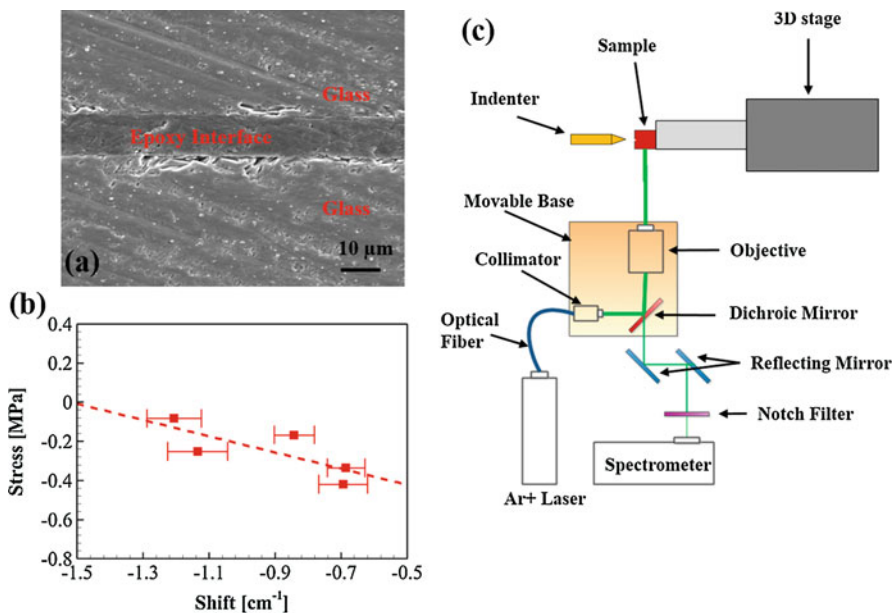


Fig. 2 (a) Image of the epoxy interface sample, (b) Raman shift-stress calibration curve, and (c) schematic of Raman measurements setup

of epoxy and the Raman shift was measured at applied loads of 100, 200, 300, 400, and 500 mN. The stresses were obtained by dividing the load by the area of the calibration sample. The measured Raman peak data was converted into shift by the equation

$$\Delta\omega = \left(\frac{1}{\lambda_{\text{laser}}} - \frac{1}{\lambda_{\text{measured}}} \right)^* 10^7 \text{ cm}^{-1}. \tag{1}$$

The change in shift was obtained by subtracting the shift at the applied load for the shift at zero load. The calibration curve for shift versus load for epoxy is given in Fig. 2b. The measurements were performed on the epoxy interface while holding the load constant. The stress distribution across interfaces is shown in Fig. 3c. The load direction is from the top of the picture. The figure shows the average values of

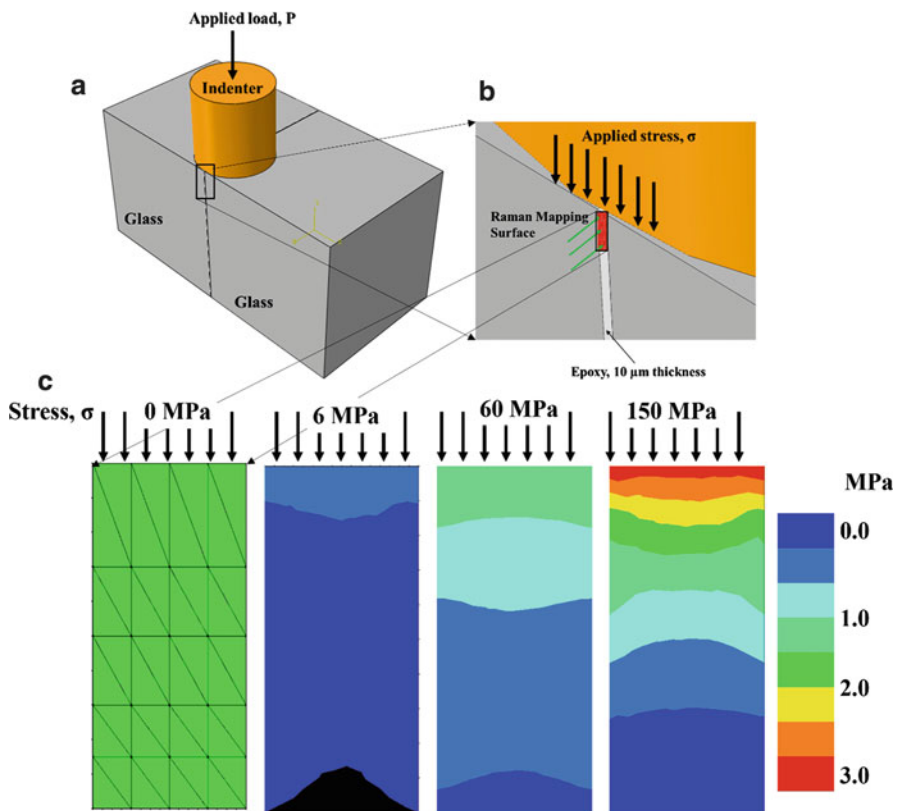


Fig. 3 (a) Schematic of epoxy interface between glasses with the indenter direction, (b) the magnified region showing the area for Raman map collection, and (c) Raman stress map on the interface with applied load of 0, 20, 200, and 500 mN. The height in each map is $80 \mu\text{m}$ and width is $25 \mu\text{m}$. Figure with 0 mN load shows the Raman data collection points

stresses as measured on the interface. This technique has been used in the literature to measure stresses tensors in the crystalline materials such as Silicon (Gan and Tomar 2014) but for the case polymers such as epoxy the average stresses can be measured (Colomban et al. 2006).

The Raman spectroscopy only provides the average stress at the interface but stress tensors in different directions are needed to fully understand the behavior of interfaces. Even in the experiments, it is difficult to measure the lateral stresses. An analytical solution is therefore developed to calculate the lateral stresses during indentation of interfaces. A schematic of the contact problem of interface is illustrated in Fig. 3. The details on the analytical solution are given in authors' previous articles.

Interface Elastic Constants

The interface elastic constants were calculated using equations in the article by Ustinov et al. (2013) for different ratios of interface modulus as compared to the main phase. The results are given in Fig. 4 for the range of ratios from 0.001 to 100. The absolute values of the interface elastic constants are also plotted to observe

Fig. 4 Absolute values of interface elastic constants calculated for different ratios of elastic modulus of main phase to interface with (a) plotting the elastic constants A11, A66, A12, A13, and (b) plotting A33 and A44

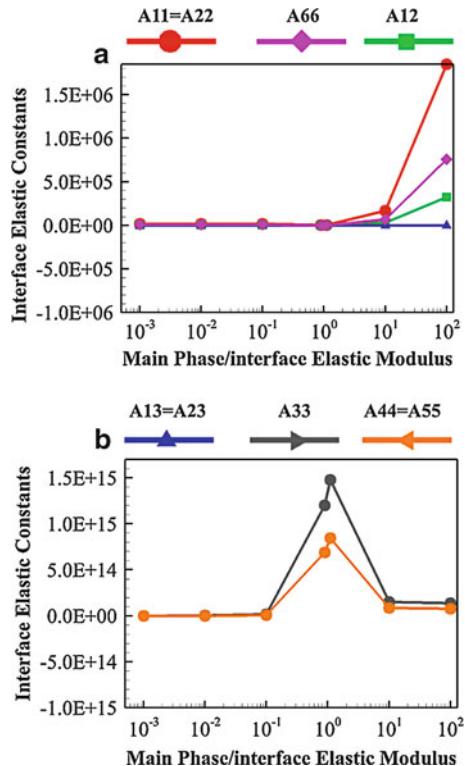


Table 1 Comparison of interface elastic constants by both methods

	Experimental solutions	Theoretical formula Ustinov et al. (2013)	Theoretical approximations Ustinov et al. (2013)
A11 (Pa-m)	34029.1	-1253260.483	53571.42857
A22 (Pa-m)	6821.673	-1253260.483	53571.42857
A33 (Pa/m)	2.03E + 14	5.74205E + 14	5.35714E + 14
A13 (Pa)	2.8E + 09	2208480565	3571428571

the variation with the change in the ratio as shown in Fig. 4. It can be seen that in the present isotropic material A11, A66, A12, A13 keeps decreasing as the modulus of both phases matches and then starts increasing again with the minimum at the point of same materials, but the opposite is true for A33 and A44 with the maximum values at the point when the interface and bulk phase has the same properties.

The interface elastic constants by both Dingreville and Qu (2008), Dingreville et al. (2014), and Ustinov et al. (2013) are calculated based on the strain energy. Per Dingreville, the interface contribution to the thermodynamic properties is defined as the excess over the values that would obtain if the bulk phases retained their properties constant up to an imaginary surface (of zero thickness) separating the two phases. The surface free (excess) energy of a near-surface atom is defined by the difference between its total energy and that of an atom deep in the interior of a large crystal. Surface free energy corresponds to the work of creating a unit area of surface, whereas surface stress is involved in computing the work in deforming a surface.

The interface elastic constants given in Table 1 are converted into surface constants based on equation from Ustinov et al. for a 10μ thickness interface. The reduced modulus measured from the indentation experiments after conversion to surface constant is also listed in the same table for comparison. The interface constants after the conversion fall under the same order as the ones calculated from the analytical relations given for interfaces.

These results show the comparison between different formulations given in the literature and a first comparison with the experimental results. The interface stress contribution is well recognized but still there is more work needed to be able to address and implement the constitutive behavior of interfaces in the engineering problems. It has a very high potential of application in areas such as interface between glass/fiber in composites, interfaces at grain boundaries in metals to mention a few.

Interface Dynamic Properties

To understand the interface mechanical properties, an idealized system with epoxy interface between two glass slides was examined under static and dynamic loading. The sample surface is shown in Fig. 2a. The impact experiments were performed

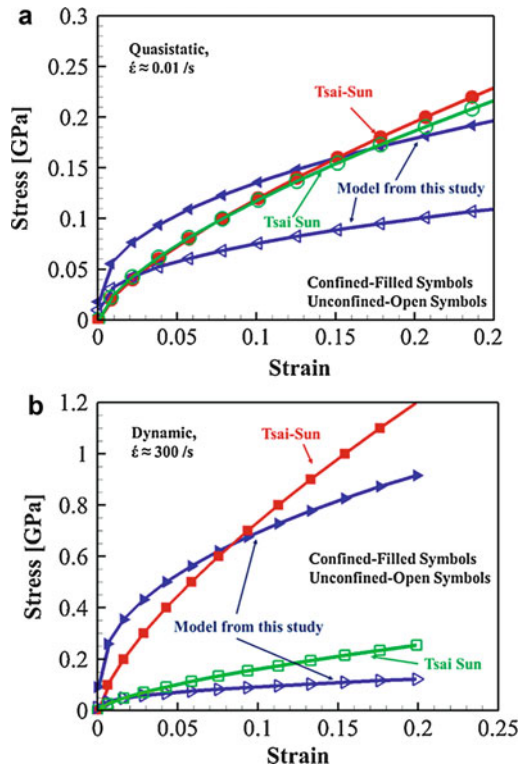
on the epoxy interface to extract the stress-strain effect of the interface at different strain rates (Verma et al. 2015). It was found that interfaces are affected by both strain rate and confinement effects during deformation. A new constitutive model is developed that couples the effect of both strain rate and lateral stresses given as

$$\sigma = (A + B\varepsilon^n) (1 + C \ln \dot{\varepsilon}^*) (1 + k\sigma_1^*) \tag{2}$$

Here, σ is the equivalent stress, ε is the equivalent plastic strain, A is the yield stress, B is the strain hardening constant, n is the strain hardening coefficient, C is the strain rate strengthening coefficient, and k is the confinement factor. $\dot{\varepsilon}^* = \frac{\dot{\varepsilon}}{\dot{\varepsilon}_{ref}}$ is the dimensionless strain rate normalized with reference strain rate, $\sigma_1^* = \frac{\sigma}{\sigma_{compressive\ strength}}$ is the dimensionless lateral stress normalized with the compressive strength of the material. The reference strain rate is taken as 1 s^{-1} and $\sigma_{compressive\ strength}$ is taken as 100 MPa. In the current experiments, the temperature was constant so the temperature effects were neglected.

The stress-strain response for the epoxy interface was analyzed at high and low strain rates as shown in Fig. 5 with the confined interface behavior labeled with closed symbols and unconfined epoxy behavior labeled with open symbols. At lower strain rates, the difference in stresses is mostly because of the confinement

Fig. 5 Fit of the current model with stress strain for (a) quasistatic loading and (b) dynamic loading



effect while in the dynamic case the strain rate effect also plays a major role. The difference in the stresses is higher in the dynamic case as evident in Fig. 5b compared to Fig. 5a. It is also compared to the stress strain behavior fitted by Tsai-Sun model on the same plot. Thus, the current model is better suited to model the behavior of materials under confined spaces such as interfaces in the composite materials, metals, ceramics, etc. A conventional way is to consider interface as a zero thickness and to not consider interface effect on the material deformation. The model in Eq. 2 takes into account the effect of interface mechanical properties on the mechanical deformation and should be considered for cases that have interface dominant geometries such as the biological materials analyzed in this work. The size effect of the interfaces is being analyzed by the authors and will be published in the future papers.

Fracture Properties of Interfaces

Cohesive Zone Model

Fracture or crack formation is the creation of new surfaces in the domain of the body. This surface creation invariably leads to loss in the global stiffness and load-bearing ability of the material, often leading to failure. Traditionally, either energy-based or stress-intensity based approaches have been employed to predict this mode of failure. The energy-based theory of failure introduced by Griffith (1921) was motivated by the inadequacy of the elastic solution that renders singular stresses at the mathematically sharp crack tip. The key idea behind the stress-intensity based theory (Irvin 1957) is the observation that the near tip crack field in isotropic linear elastic materials is similar for all specimen geometry and loading conditions, to within a constant.

To address the inconsistency of infinite stresses at the crack-tip, a theory involving a process zone, equivalent to a plastic zone in elastic-plastic fracture mechanics, was presented by Barenblatt (1962) and Dugdale (1960), for crack propagation in homogeneous isotropic materials. The main assumption was that in the vicinity of crack tip, opposing faces of the crack are bounded by the molecular cohesive forces. As the body is loaded, the two faces undergo significant deformation, which eventually leads to loss of interatomic cohesion and traction free surface creation. The cohesive forces are concentrated near a small, but finite region of the continuum crack tip and drop to zero within few atomic distances from the crack tip. This idea of the cohesive traction applied in a small cohesive zone removes the difficulty of crack tip singularity in LEFM and is called Cohesive Zone Model (CZM). A cohesive zone model assumes a relation between the normal (and shear) traction and the opening (and sliding) displacement, and is capable of capturing the debonding process of interfaces.

There are several numerical studies of failure based on the cohesive zone models (Barua et al. 2012a, b, 2013a, b; Barua and Zhou 2011; Kim et al. 2014; Tvergaard 2003; Tvergaard and Hutchinson 1992, 1993; Camacho and Ortiz 1996; Zhong

1999; Zhong and Knauss 2000, 1997; Zhong and Meguid 1997; Needleman and Xu 1993; Kubair et al. 2002a, b; Samudrala et al. 2002; Samudrala and Rosakis 2003). Using Rose et al.'s (1981) cohesive law that is primarily for bimetallic interfaces, Levy (1994) studied the debonding of a circular fiber in an infinite matrix subject to equibiaxial load. There are also analytical studies of interface debonding based on cohesive zone models, and these studies focus on the effect of interface debonding on the macroscopic behavior of composite materials. Almost all analytic studies are limited to the linear cohesive law, i.e., a linear relation between the normal (and shear) traction(s) and the opening (and sliding) displacement(s) of interfaces (!!! INVALID CITATION !!!) n.d.; Hashin 1991a, b, 2002; Qu 1993; Wu et al. 1999). There are very few analytic studies on the effect of nonlinear cohesive law (Levy 1996, 2000).

Experimental observations show distinctive characteristics in micromechanical failure mechanisms in peel and shear fracture, thus the cohesive behavior is expected to be mode dependent (Chai 2004, 2003; Chai and Chiang 1996; Roy et al. 1999). A frequently used coupled cohesive law is developed by Tvergaard and Hutchinson (1993), using a dimensionless separation parameter. A drawback with this model is that the fracture energy is the same in all mode mixities (Högborg 2006). Another theoretically motivated exponential cohesive law is developed by Xu and Needleman (Needleman and Xu 1993). Van den Bosch et al. (2006) showed that this exponential formulation only realistically describes mixed mode behavior in a special case: The fracture energy is equal in normal and tangential direction. Park and Paulino (2013) have extensively reviewed a number of CZMs and concluded that the available models should be used with great caution for mix-mode cohesive fracture.

There are several CZMs that have been suggested in literature (Park and Paulino 2013; Needleman 2014; Shet and Chandra 2004). The constitutive behavior of the cohesive model is formulated as a traction-separation law (TSL), which relates the traction, T , to the separation, Δ , the latter representing the displacement jump across the interface. These cohesive parameters can be extracted from load-displacement curve obtained from a simple tension or shear test (Zhu et al. 2009). One approach is to use a J-integral versus end opening curve to get the cohesive strength and the critical separation. Molecular dynamics (MD) simulations have also been used previously to obtain cohesive zone parameters of the interfaces (Dandekar and Shin 2011; Zhou et al. 2008, 2009; Yamakov et al. 2006). Stress-displacement curve obtained from MD simulation is used as a traction-separation curve and the maximum stress is used as cohesive strength of the interface. The total area under stress-separation curve is taken as the cohesive fracture energy.

Interface Thickness Effect

In this section we will show the effect of interface thickness on the fracture properties by studying the effect of grain boundary thickness in a polycrystalline material. In a typical polycrystalline metal, GBs are significantly thinner (~ 10 nm)

than the grains ($\sim\mu\text{m}$). The accurate prediction of crack propagation through GBs and interfaces while simultaneously predicting crack propagation through grains can be difficult due to the factors that include: significant difference in the length scales of GBs and grains, respectively; unknown GB strength properties; embrittlement effects of GBs owing to GB chemistry, etc. Considering the fact that Ni additions are likely to be segregated along GBs in polycrystalline W, (Gupta et al. 2007), change in GB mechanical property plays a significant role in enhancing or degrading of the materials failure resistance.

The GB simulation model is constructed based on the image of a realistic Ni doped W GB from a HRTEM image (Gupta et al. 2007). Fully saturated Ni-doped W GB has a thickness of about 0.6 nm (Gupta et al. 2007). However, when calculating the GB properties one must insert a few atoms on either side of GB, Fig. 1, resulting in an interface structure.

A 3-D GB interface model of Ni-doped W GB, Fig. 6d, is developed for (Prakash et al. 2016) fracture simulations using a combination of XFEM with the cohesive finite element method (CFEM) in order to characterize the influence of GB thickness on fracture toughness. For the constitutive description of W grains and Ni doped W GBs in a polycrystalline W microstructure, a general elastic-plastic material model is used with parameters derived based on the data presented in the ab-initio modeling work of Lee and Tomar (2013, 2014). Cohesive parameters were evaluated from the stress-displacement curve obtained from the ab-initio calculation.

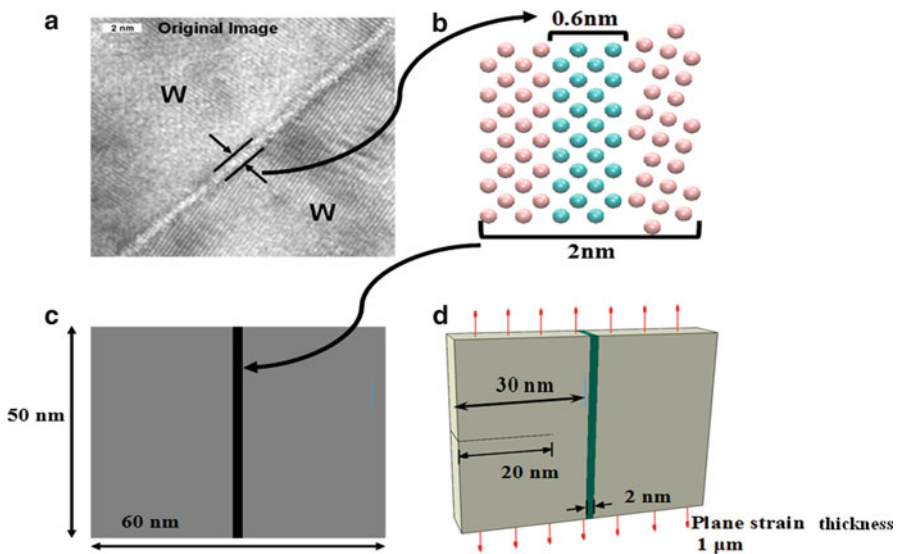


Fig. 6 A schematic showing (a) HRTEM image of polycrystalline Ni doped W GB (Gupta et al. 2007), (b) atomic structure of W GB analyzed by Lee and Tomar (2013, 2014), (c) 2-D, and (d) 3-D image of the continuum GB element analyzed for crack propagation

Although the total GB volume is relatively small, the thickness of GBs can greatly affect crack propagation. Lee and Tomar (2013, 2014) found that the minimum thickness of such an interphase that contains GB and few atomic layers around GB should be at least 2 nm in order to predict thickness independent properties of GB in *ab initio* simulations. In order to understand the influence of GB thickness on simulation results, crack propagation through interfaces with GB thickness of 2 nm, 4 nm, and 8 nm was analyzed. As shown in Fig. 7, change in thickness of GB does not significantly influence either the crack tip position change or crack tip elastic strain energy release (calculated based on elastic strains surrounding crack tip) indicating that GB thickness change at the length scale of

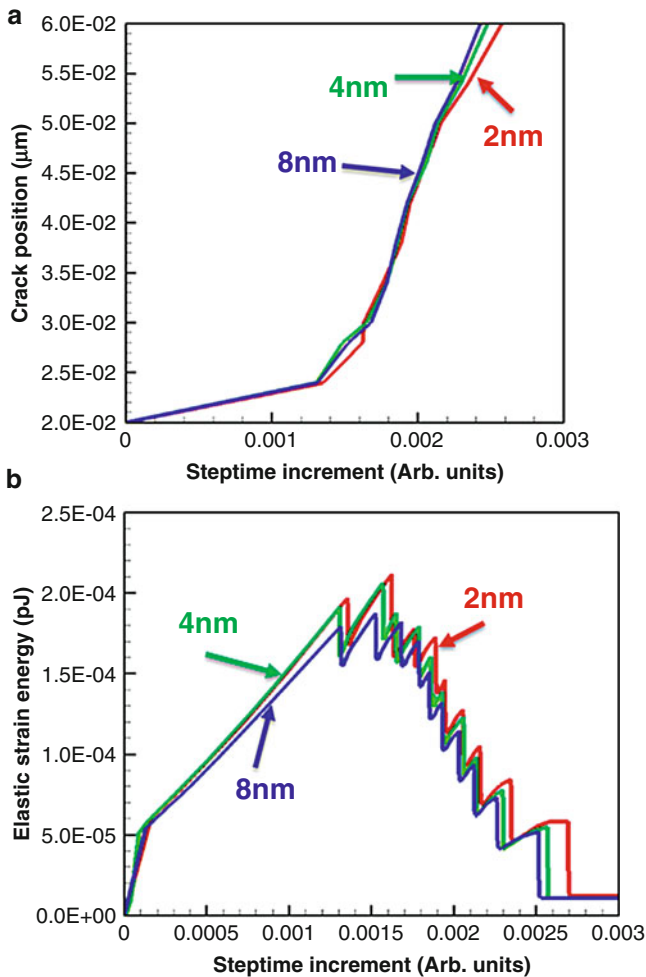


Fig. 7 Effect of GB thickness on (a) crack tip position and (b) strain energy release with respect to the time-step increment

materials model that corresponds to atomic scale does not affect interface crack propagation characteristics.

In order to understand the effect of length scale, as discussed earlier, scaled up interface structured with different GB thickness are analyzed for crack propagation. Figure 7 shows the total plastic energy dissipation per unit volume and crack tip strain energy change per unit volume change as interface specimen is loaded and crack propagation undergoes. The plastic energy dissipation is calculated based on plastic strain in the region surrounding the crack tip. As shown, the effect of length scale on fundamental energetic quantities related to crack propagation is insignificant (Fig. 8).

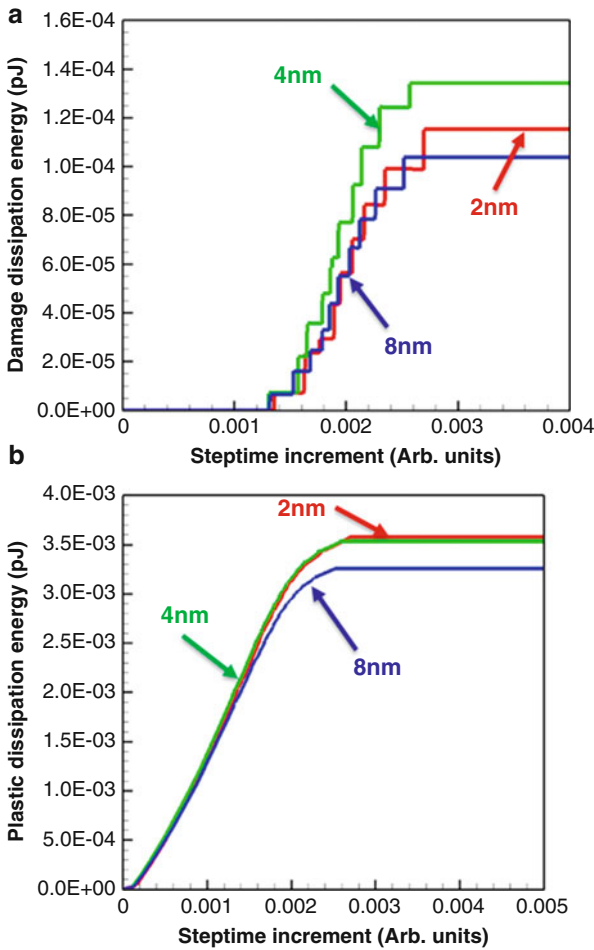
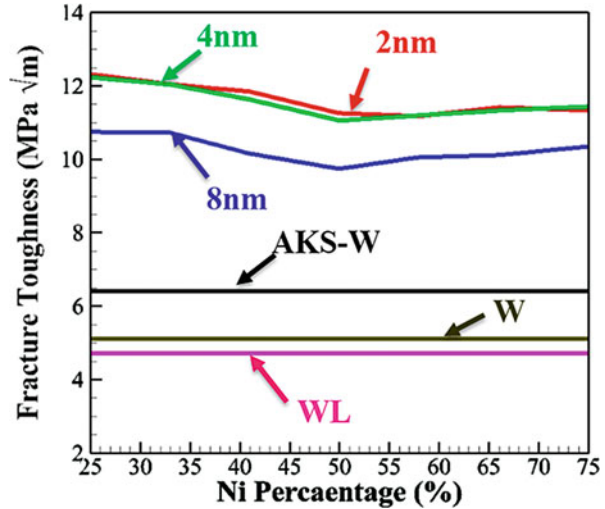


Fig. 8 Effect of GB thickness on the (a) damage dissipation energy and (b) plastic dissipation energy of the 3-D model

Fig. 9 Fracture toughness with respect to Ni percentage in GB and a comparison of fracture toughness of W interface specimens at the three scales examined and experimentally available values for polycrystalline W (Gludovatz et al. 2010)



As shown in Fig. 9, a clear scale dependence of fracture toughness is seen with the value corresponding to 100 nm thick interface length scale approaching experimental values. Earlier, Mai and Lawn (1987) have used theoretical fracture mechanics-based approach to calculate scale-dependent fracture toughness (Mai and Lawn 1987).

Theoretical calculations showed that the fracture toughness has an intrinsic length-scale dependence, which matches with the result in the presented study. In the work of Mai and Lawn (1987), the range of fracture toughness was found to be 2–7 MPa for Al, and it has been observed that the fracture toughness tends to decrease significantly as length-scale of system becomes smaller, similar to what is observed here. This result establishes the length scale dependence of fracture toughness in examined interfaces and does not refute the use of continuum elements at the small length scale of 2 nm.

Conclusions

A new technique based on nanomechanical Raman spectroscopy is presented to calculate the interface stresses and elastic constants using an analytical model. The elastic constants are then compared with the strain energy frameworks provided in the literature. The comparison between both methods shows the dependence of the interface elastic constants on the thickness of the interface. The elastic constants calculated from the stress-strain data matches the theoretical values after the thickness effect correction.

The interfaces are believed to play an important role in the observed behavior. The role of interfaces is identified by performing experiments on an idealized system of glass-epoxy interfaces. The confinement effect on the interfaces along

with the effect of strain-rate was found to play a major role in the deformation of the examined interfaces. A new model capturing both strain rate and confinement effects is developed for strain rates up to 10^3 s^{-1} in this paper to account for confinement effect and strain rate effect coupling.

Interface thickness effects were studied for the Ni-doped Tungsten polycrystalline material. For small length scale, as in the case of GB, it has been shown that the fracture toughness at the interface is affected by the GB thickness. However, the crack tip stress and the strain energy at the interface are not affected. Further investigation on the effect of interface thickness at large length scale is being studied by the authors.

References

- G.I. Barenblatt, *The Mathematical Theory of Equilibrium Cracks Formed in Brittle Fracture* (Armed Services Technical Information Agency, Arlington, 1962)
- A. Barua, Y. Horie, M. Zhou, *J. Appl. Phys.* **111**, 054902 (2012a)
- A. Barua, Y. Horie, M. Zhou, *Proceedings of the royal society a: mathematical. Phys. Eng. Sci.* **468**, 3725 (2012b)
- A. Barua, S. Kim, Y. Horie, M. Zhou, *J. Appl. Phys.* **113**, 184907 (2013a)
- A. Barua, S.P. Kim, Y. Horie, M. Zhou, *Mater. Sci. Forum* **767**, 13 (2013b)
- A. Barua, M. Zhou, *Model. Simul. Mater. Sci. Eng.* **19**, 055001 (2011)
- M.J. van den Bosch, P.J.G. Schreurs, M.G.D. Geers, *Eng. Fract. Mech.* **73**, 1220 (2006)
- F. Boßelmann, P. Romano, H. Fabritius, D. Raabe, M. Epple, *Thermochim. Acta* **463**, 65 (2007)
- G.T. Camacho, M. Ortiz, *Int. J. Solids Struct.* **33**, 2899 (1996)
- H. Chai, *Int. J. Solids Struct.* **40**, 6023 (2003)
- H. Chai, *Int. J. Fract.* **130**, 497 (2004)
- H. Chai, M.Y.M. Chiang, *J. Mech. Phys. Solids* **44**, 1669 (1996)
- P.-Y. Chen, A.Y.-M. Lin, J. McKittrick, M.A. Meyers, *Acta Biomater.* **4**, 587 (2008)
- P. Colomban, G. Gouadec, J. Mathez, J. Tschember, P. Pérès, *Compos. A: Appl. Sci. Manuf.* **37**, 646 (2006)
- C.R. Dandekar, Y.C. Shin, *Compos. A: Appl. Sci. Manuf.* **42**, 355 (2011)
- R. Dingreville, A. Hallil, S. Berbenni, *J. Mech. Phys. Solids* **72**, 40 (2014)
- R. Dingreville, J. Qu, *J. Mech. Phys. Solids* **56**, 1944 (2008)
- D.S. Dugdale, *J. Mech. Phys. Solids* **8**, 100 (1960)
- E. Flores-Johnson, L. Shen, I. Guiamatsia, G.D. Nguyen, *Compos. Sci. Technol.* **96**, 13 (2014)
- M. Gan, V. Tomar, *Rev. Sci. Instrum.* **85**, 013902 (2014)
- B. Gludovatz, S. Wurster, A. Hoffmann, R. Phipps, *Int. J. Refract. Met. Hard Mater.* **28**, 674 (2010)
- A.A. Griffith, *Philosophical transactions of the royal society a: mathematical. Phys. Eng. Sci.* **221**, 163 (1921)
- V.K. Gupta, D.-H. Yoon, H.M. Meyer Iii, J. Luo, *Acta Mater.* **55**, 3131 (2007)
- Z. Hashin, *J. Mech. Phys. Solids* **39**, 745 (1991a)
- Z. Hashin, *Trans. ASME* **58**, 444 (1991b)
- Z. Hashin, *J. Mech. Phys. Solids* **50**, 2509 (2002)
- J.L. Högberg, *Int. J. Fract.* **141**, 549 (2006)
- N. Pagano and G.P. Tandon: *Mechanics of Materials*, **9**, 49 (1990)
- G.R. Irwin, *J. Appl. Mech.* **24**, 361 (1957)
- S. Kim, A. Barua, Y. Horie, M. Zhou, *J. Appl. Phys.* **115**, 174902 (2014)
- D.V. Kubair, P.H. Geubelle, Y.Y. Huang, *Eng. Fract. Mech.* **70**, 685 (2002a)
- D.V. Kubair, P.H. Geubelle, Y.Y. Huang, *J. Mech. Phys. Solids* **50**, 1547 (2002b)
- H. Lee, V. Tomar, *Comput. Mater. Sci.* **77**, 131 (2013)

- H. Lee, V. Tomar, *Int. J. Plast.* **53**, 135 (2014)
- A. Levy, *J. Mech. Phys. Solids* **42**, 1087 (1994)
- A. Levy, *J. Appl. Mech.* **63**, 357 (1996)
- A. Levy, *J. Appl. Mech.* **67**, 727 (2000)
- Y.-W. Mai, B.R. Lawn, *J. Am. Ceram. Soc.* **70**, 289 (1987)
- G. Mayer, *J. Mech. Behav. Biomed. Mater.* **4**, 670 (2011)
- A. Needleman, *Procedia IUTAM* **10**, 221 (2014)
- A. Needleman, X.P. Xu, *Model. Simul. Mater. Sci. Eng.* **1**, 111 (1993)
- K. Park, G.H. Paulino, *Appl. Mech. Rev.* **64**, 060802 (2013)
- C. Prakash, H. Lee, M. Alucozai, V. Tomar, *Int. J. Fract.* **199**, 1 (2016)
- J. Qu, *Mech. Mater.* **14**, 269 (1993)
- T. Qu, D. Verma, M. Alucozai, V. Tomar, *Acta Biomater.* **25**, 325 (2015)
- T. Qu, D. Verma, M. Shahidi, B. Pichler, C. Hellmich, V. Tomar, *MRS Bull.* **40**, 349 (2015)
- D. Raabe, P. Romano, C. Sachs, H. Fabritius, A. Al-Sawalmih, S.-B. Yi, G. Servos, H. Hartwig, *Mater. Sci. Eng. A* **421**, 143 (2006)
- J.H. Rose, J. Ferrante, J.R. Smith, *Am. Phys. Soc.* **47**, 675 (1981)
- A.Y. Roy, R. Narashimhan, P.R. Arora, *Acta Mater.* **47**, 1587 (1999)
- O. Samudrala, Y. Huang, A.J. Rosakis, *J. Mech. Phys. Solids* **50**, 1231 (2002)
- O. Samudrala, A.J. Rosakis, *Eng. Fract. Mech.* **70**, 309 (2003)
- C. Shet, N. Chandra, *Mech. Adv. Mater. Struct.* **11**, 249 (2004)
- V. Tvergaard, *Eng. Fract. Mech.* **70**, 1859 (2003)
- V. Tvergaard, J.W. Hutchinson, *J. Mech. Phys. Solids* **40**, 1377 (1992)
- V. Tvergaard, J.W. Hutchinson, *J. Mech. Phys. Solids* **41**, 1119 (1993)
- K.B. Ustinov, R.V. Goldstein, V.A. Gorodtsov, On the modeling of surface and interface elastic effects in case of eigenstrains, in *Surface Effects in Solid Mechanics: Models, Simulations and Applications*, ed. by H. Altenbach, F. N. Morozov (Springer Berlin Heidelberg, Berlin, 2013), p. 167
- D. Verma, J. Singh, A.H. Varma, V. Tomar, *JOM* **67**, 1694 (2015)
- D. Verma, V. Tomar, *Mater Sci Eng C Mater Biol Appl* **44**, 371 (2014a)
- D. Verma, V. Tomar, *J Bionic Eng* **11**, 360 (2014b)
- D. Verma, V. Tomar, *J. Mater. Res.* **30**, 1110 (2015a)
- D. Verma, V. Tomar, *Mater Sci Eng C Mater Biol Appl* **49**, 243 (2015b)
- Y. Wu, Z. Ling, Z. Dong, *Int. J. Solids Struct.* **37**, 1275 (1999)
- V. Yamakov, E. Saether, D.R. Phillips, E.H. Glaessgen, *J. Mech. Phys. Solids* **54**, 1899 (2006)
- X.A. Zhong, *Mech. Adv. Mater. Struct.* **6**, 1 (1999)
- X.A. Zhong, W.G. Knauss, *Trans. ASME* **119**, 198 (1997)
- X.A. Zhong, W. Knauss, *Mech. Adv. Mater. Struct.* **7**, 35 (2000)
- Z. Zhong, S.A. Meguid, *J. Elast.* **46**, 91 (1997)
- X.W. Zhou, N.R. Moody, R.E. Jones, J.A. Zimmerman, E.D. Reedy, *Acta Mater.* **57**, 4671 (2009)
- X.W. Zhou, J.A. Zimmerman, E.D. Reedy, N.R. Moody, *Mech. Mater.* **40**, 832 (2008)
- Y. Zhu, K.M. Liechti, K. Ravi-Chandar, *Int. J. Solids Struct.* **46**, 31 (2009)



Nanostructural Response to Plastic Deformation in Glassy Polymers

12

George Z. Voyiadjis and Aref Samadi-Dooki

Contents

Introduction	378
Theory	381
Free Volume Evolution	381
STZ Nucleation Energy Evolution	384
Results and Discussion	388
Effect of the Rate	389
Effect of the Temperature	391
Effect of the Thermal History	393
Implication of the Model for Shear Banding and Indentation Size Effect in PGs	395
Concluding Remarks	397
References	397

Abstract

A closed form stress-strain relation is proposed for modeling the postyield behavior of amorphous polymers based on the shear transformation zones (STZs) dynamics and free volume evolution. Use is made of the classical free volume theory by Cohn and Turnbull (J Chem Phys 31:1164, 1959), and also STZ-mediated plasticity model for amorphous metals by Spaepen (Acta Metall 25:407, 1977) and Argon (Acta Metall 27:47, 1979) for developing a new homogenous plasticity framework for glassy polymers. The variations

G. Z. Voyiadjis (✉)

Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge, LA, USA

e-mail: voyiadjis@eng.lsu.edu

A. Samadi-Dooki

Computational Solid Mechanics Laboratory, Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge, LA, USA

e-mail: asamad3@lsu.edu

of free volume content and STZs activation energy during large deformation are parametrized considering the previous experimental measurements using positron annihilation lifetime spectroscopy (PALS) and thermal analysis with differential scanning calorimetry (DSC), respectively. The proposed model captures the softening-hardening behavior of glassy polymers at large strains with a single formula. This study shows that the postyield softening of the glassy polymers is a result of the reduction of the STZs nucleation energy as a consequence of increased free volume content during the plastic straining up to a steady-state point. Beyond the steady-state strain where the STZ nucleation energy reaches a plateau, the increased number density of STZs, which is required for finite strain, brings about the secondary hardening continuing up to the fracture point. This model also accurately predicts the effect of strain rate, temperature, and thermal history of the sample on its postyield behavior which is in consonance with experimental observations. Implication of the model for interpreting the localization and indentation size effect of polymers is also discussed.

Keywords

Plasticity · Polymer · Amorphous · Free volume · Shear transformation · Stress · Strain · Energy · Deformation · Microstructure · Chain · Rate · Glass

Introduction

Many industries have turned their attention to polymers as the materials that can be used as key structural elements with excellent compatibilities with the efficiency-centered design criterion. Polymers are light-weight, inexpensive, transparent, and extremely formable, and with some minor enhancements (like reinforcing with fibers in polymer matrix composites) can be utilized as load-bearing elements under severe loading conditions. Polymeric glasses (PGs), which consist the majority of polymers, are amorphous materials with no arrangement in their microstructure. From the mechanical point of view, while the preyield behavior of this type of materials can be considered as viscoelastic, they show a very distinct deformation response when loaded beyond their yield point (Argon 2013; Boyce et al. 1988). Of the characteristic behavior of glassy polymers within the postyield region is the considerable primary softening which takes place right after yielding and continues up to a certain steady state strain, and a secondary hardening which extends up to the fracture point as demonstrated in Fig. 1.

Many efforts have been made to unfold and quantitatively describe the aforementioned unique postyield behavior of PGs during the past one-half century. The earliest significant model for the plastic deformation of PGs is the one proposed by Eyring (1936). With considering the molecular rearrangement as the key factor mediating the plastic deformation, Eyring (and later Ree and Eyring 1955) suggested that the restriction or activation of the spatial movement of the chains with altering the loading rate and/or the temperature is responsible for elevation or decreasing of the PGs' resistance to deformation. With considering

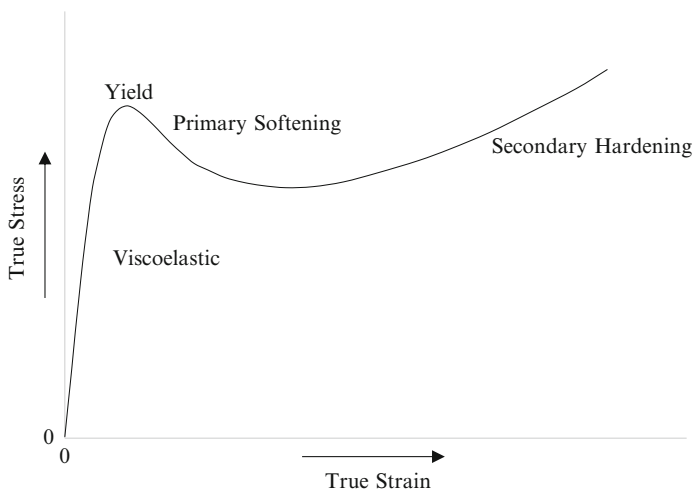


Fig. 1 A schematic representation of the generic uniaxial stress-strain behavior of glassy polymers (Reproduced from Voyiadjis and Samadi-Dooki 2016 with permission from AIP)

the chain deformation as the main contributing factor to the strain resistance of the macromolecular systems, Robertson (1966) proposed a model for the plastic deformation of these materials which was different from the rubber-like deformation mechanism. Robertson's model, however, is able to correctly predict the large deformation behavior of PGs at temperatures close to the glass transition temperature (T_g) only. Later, Argon (1973) came up with a model which considered the localized antisymmetric twists along the macromolecular chains as the mechanism for the permanent deformation. This model was formulated by considering the energy of the twisting mechanism to be calculated similar to that of a pair of adjacent disclinations. Known as "double kink" model, Argon's formalism successfully predicts the rate, temperature, and pressure dependence of the yielding behavior of amorphous polymers. Hence, it was extended by others for modeling the postyield plastic deformation (Arruda and Boyce 1991; Arruda et al. 1993, 1995; Boyce et al. 1988) and strain gradient plasticity (Lam and Chong 1999; Voyiadjis et al. 2014) of PGs. Among the most noticeable developments to Argon's model is the one proposed by Boyce et al. (1988) who used this theory to express the resistance of the macromolecules to segmental rotation. Accordingly, this model suggests that once this resistance is overcome, the chain entanglement opening causes a considerable softening in stress-strain behavior of the PGs. At higher strains, the chain alignment is considered to be the dominant factor which results in secondary hardening of the material behavior.

For crystalline materials with defined microstructural order, the primary carriers of plastic deformation are nanostructural defects like dislocations and point defects (Argon 2008; Shodja et al. 2013; Voyiadjis and Yaghoobi 2015). In contrast, the plastic deformation of amorphous polymers has traditionally been thought as the conformational changes in macromolecular systems, like chain rearrangements.

Since no preferred long-range order exists in the intra- and intermolecular space these amorphous solids, considering the chain rearrangement as the sole mechanism responsible for the permanent deformation of this type of solids, seem to be reasonable. Accordingly, it is expected that the stored energy in the permanently deformed polymers to be small which can be justified by conformational rearrangement. However, the calorimetric analyses by Oleinik and coworkers (Oleinik 1991; Oleinik et al. 2007) for a number of amorphous polymers revealed a large amount of stored energy during the course of plastic deformation which can never be attributed to the chain rearrangement mechanism. Instead, Oleinik (1991) suggested that this large amount of observed stored energy is due to the elastic field around some defect sites which form during the plastic deformation, like Eshelby's inclusion model (Eshelby 1957; Mura 1987). These defect sites are, in fact, the localized permanent deformation inclusions which encompass slip, glide, and shear rotation of macromolecular chains. The formation of these inclusions is mostly enforced by the shear component of the applied stress (Oleinik 1991), hence, they are termed shear transformation zones (STZs). It is worth noting that STZs were originally proposed to be responsible for the plastic deformation of metallic glasses (MGs) (Argon 1979; Spaepen 1977), but they are now believed to be the main mechanism for the plastic deformation of all types of disordered solids (Argon 2013). Unfortunately, the direct observation of STZs is not practical at this time, because they are localized transition zones rather than being actual defects (Wang et al. 2013); nevertheless, physical evidences exist for proving the nucleation of these sites during the inelastic deformation of solid state glasses (Falk 2007; Oleinik 1991; Pauly et al. 2010).

From the shear transformation plasticity point of view, the excessive applied stress triggers the nucleation of STZs, and the kinetics of inelastic deformation is determined by the rate of the formation of these sites (Argon 1979; Spaepen 1977). Hence, the dynamics of the STZs seems to be a determining factor for analyzing the plastic deformation of glassy state solids. For the case of MGs, this hypothesis has become the principal approach to study their yielding and postyield behavior. Most noticeably, Langer and Pechenik (2003) have proposed that the rate of the STZs' annihilation and creation needs to be proportional to the rate of the energy dissipation during plastic deformation in order to satisfy the thermodynamic consistency with considering the first and second laws. Since this model was successful in qualitatively capturing the inelastic deformation behavior of MGs, it was extended by Falk et al. (2004) with considering the role of thermal fluctuations on the atomic rearrangement during large deformation of MGs.

From the microscopic point of view, the localized deformation sites tend to form around the "weak" areas in the bulk of the solids. Amorphous solids possess a considerable amount of "voids" in their structure, the presence of which is dictated by the very nature of this type of materials as disordered solids (Flores et al. 2002; Hofmann et al. 2002; Jean et al. 2013; Utracki and Jamieson 2011; Yavari et al. 2005). While these voids can serve as weak points for STZs formation, existence of a correlation between the evolution of STZs and that of the free volume sites

which is observed during plastic deformation of amorphous materials like MGs (de Hey et al. 1997) seems to be reasonable. Accordingly, for the case of MGs, some researchers developed models which consider the adjustment of STZs nucleation energy with the variation of free volume content as a state variable, similar to the one proposed by Li et al. (2013).

Although the shear transformation-mediated plasticity is the accepted inelastic deformation mechanism for all of the amorphous solids (Argon 2013), the number of studies using this approach for PGs is limited. Among these studies, the works by Hasan and coworkers notably consider the evolution of the number density of STZs during the large deformation as a determining factor for plastic deformation (Hasan and Boyce 1995; Hasan et al. 1993). This model is capable of capturing the effect of the rate, temperature, and thermal history of the sample on the postyield softening behavior of PGs. Argon (2013) also proposed a relation for shear transformation-mediated plasticity of PGs based on his molecular dynamics simulations of amorphous silicon (Argon and Demkowicz 2006, 2008) to capture the effect of the thermal history of the sample on PGs' overall postyield softening behavior. With defining the liquid-like sites as the fertile sites for shear transformation, Argon's model describes the evolution of these liquid-like densities during the plastic deformation.

In this chapter, a new model is presented for the postyield deformation of PGs which formulates the competition between STZs' nucleation energy variation due to the evolution of the free volume content during large deformation and the increase of STZs' number density to bring about the permanent plastic deformation. The variation of the number density of the STZs is directly related to the plastic strain, and the variation of the deformation energy and free volume are quantified based on the previous experimental studies by Oleinik (1991) and Hasan et al. (1993), respectively. Accordingly, effects of the temperature, strain rate, and thermal history of the polymer on its plastic deformation are investigated. With incorporating the microgeometrical and micromechanical properties of the STZs of Poly(methyl methacrylate) (PMMA) obtained from nanoindentation experiments and presented in the authors previous study (Malekmotiei et al. 2015), the results based on the proposed model are obtained and compared with those of previous experimental observations for uniaxial compressive deformation of this PG.

Theory

Free Volume Evolution

In developing the current theory, the plastic deformation is assumed to be a homogeneous flow which is mainly mediated by the nucleation and accumulation of shear events rather than a sudden development of shear bands. Accordingly, the kinetics relation which presents the relation between the plastic shear strain rate

$\dot{\gamma}^p$ and the shear stress in the plastic state τ^p can be used which was originally developed by Spaepen (1977) and Argon (1979) as:

$$\dot{\gamma}^p = \dot{\gamma}_0^p c_f \exp\left(\frac{-\Delta F}{k_B T}\right) \sinh\left(\frac{\gamma^T \Omega \tau^p}{2k_B T}\right), \quad (1)$$

in which $\dot{\gamma}_0^p$ is a preexponential constant, c_f represents the fraction of the potential jump sites, ΔF is the nucleation energy barrier, γ^T is the transformation strain, Ω is the volume of an STZ, T is the absolute temperature, and k_B is the Boltzmann constant.

For the case of MGs, potential jump sites are the interatomic spaces with excessive free volume which can act as sources for atomic jumps that result in STZs formation. The fraction of potential jump sites in an amorphous material can be expressed by the relation developed by Cohen and Turnbull (1959) based on their probability analysis as:

$$c_f = \exp\left(\frac{-\lambda v^*}{v_f}\right), \quad (2)$$

where λ is a geometrical factor which is in the range of 1/2 to 1, v^* is the minimum size of an excessive free volume which can serve as a jump site, and v_f is the average free volume per atom. Assuming an even distribution of free volume in the bulk of the solid, the average free volume per atom may be defined as:

$$v_f = \frac{V_f}{N}, \quad (3)$$

in which V_f is the total free volume content of the sample and N is its total number of atoms.

Despite MGs which have atomic microstructures, PGs are mainly composed of huge covalently bonded molecules. Hence, the “atomic jump” concept seems to require a revisit and redefinition for STZ-mediated plasticity of polymers. Glassy polymers possess an average free volume in the order of 5% with the average size of about a couple of hundred cubic Angstroms (Jean 1990). This size is about the size of a handful of monomers of common glassy polymers. On the other hand, each STZ in PGs is calculated to encompass a few thousands of monomers (Argon 2013; Malekmotiei et al. 2015; Samadi-Dooki et al. 2016). Although the size of the free volume voids seems to be small compared to that of the STZs, these voids can be assumed to provide weak spots for nucleation of STZs. The fact that free volume content increases with increasing the STZs number density in PGs (Hasan et al. 1993) suggests a reciprocal effect between the evolution of these two parameters in polymers. Hence, it could be reasonably assumed that free volume is redistributed at the STZs’ sites upon their formation. Accordingly, N in Eq. 3 might be referred to as the number of STZs in PG, and the total free volume can be assumed to be a fraction of the total plastic volume V_p as:

$$V_f = f V_p, \quad (4)$$

where f stands for the fraction of the free volume.

Several experimental techniques have been used for evaluating the free volume content of polymers. The positron annihilation lifetime spectroscopy (PALS) is one of them which offers a reliable method for measuring the concentration and average size of the free volume holes in the bulk of the polymeric glass (Jean et al. 2013). In this method, positrons are injected into the bulk of the polymer from a radioactive source. Upon injection, part of the positrons interact with the electrons and form a particle called positronium (Ps), which can exist in form of para-positronium (p-Ps) with antiparallel spins and ortho-positronium (o-Ps) with parallel spins. Among these two particles, the transitional properties of o-Ps have important implications on the free volume of polymers. In particular, the intensity of o-Ps particles represents the concentration of the free volume sites, and its lifetime correlates with the average radius, R , of these voids as (Jean 1990):

$$\frac{1}{\tau_{o-Ps}} = 2 \left[1 - \frac{R}{R + \Delta R} + \frac{1}{2\pi} \sin \left(\frac{2\pi R}{R + \Delta R} \right) \right], \quad (5)$$

in which $\Delta R = 1.656 \text{ \AA}$ is an empirical constant and τ_{o-Ps} is the o-Ps lifetime in nanoseconds. The free volume content of the polymer can be related to the average volume and concentration of these sites as (Jean 1990; Wang et al. 1990):

$$f = \frac{4}{3} \pi R^3 (A I_{o-Ps} + B), \quad (6)$$

with I_{o-Ps} representing the o-Ps concentration and $A = 0.018$ and $B = 0.39$ are the experimentally found constants. The two o-Ps pertaining parameters, τ_{o-Ps} and I_{o-Ps} , are both temperature dependent; however, only the o-Ps annihilation lifetime τ_{o-Ps} is dependent on the plastic strain and thermal history of the polymer (Hasan et al. 1993). Accordingly, the free volume content of the polymer can be taken as a function of temperature T , shear plastic strain γ^p , and thermal history of the sample θ as:

$$f \equiv f(\gamma^p, T, \theta). \quad (7)$$

To come up with a relation for the fraction of the jump sites of Eq. 2 which is fully defined based on the measurable parameters, one needs to present alternate definitions for the number of STZs, N , and the factor λv^* . Since it presents a factor for the minimum size of the free volume that can accommodate an STZ, the factor λv^* can be assumed to be proportional to the shear activation volume V^* (Li et al. 2013), which itself is defined as:

$$V^* = \alpha \gamma^T \Omega, \quad (8)$$

in which α is a factor for considering the dilatation effect and is close to unity for PGs (Ho et al. 2003; Malekmotiei et al. 2015; Samadi-Dooki et al. 2016). On the other hand, the number of the shear transformation zones can be assumed to be proportional to the total shear plastic strain (Argon 2013) and may be calculated as:

$$N \propto \frac{\gamma^p V_p}{\gamma^T \Omega}, \quad (9)$$

the parameters in which have already been defined. Finally, by incorporating Eqs. 3, 4, 7, 8, and 9, the fraction of the potential jump sites of Eq. 2 can be rearranged as:

$$c_f = \exp\left(-\frac{k\gamma^p}{f(\gamma^p, T, \theta)}\right), \quad (10)$$

in which k is a single constant that accounts for the various proportionality factors. Apparently, this equation presents the fractional free volume as a function of measurable variables such as temperature, plastic strain, and thermal history of the sample. While plastic strain is a dependent variable which is analytically present in Eq. 10, the variation of f with temperature, plastic strain, and thermal history of the sample is extracted from appropriate experimental studies using PALS (Hasan et al. 1993; Hristov et al. 1996).

STZ Nucleation Energy Evolution

As it was mentioned before, the STZs formed within the bulk of the glassy materials can be treated as nonelastic inclusions. Eshelby (1957) proposed a rigorous mathematical model for finding the elastic fields inside and outside the inclusions which has been used by many researchers as a homogenization technique in various types of problems (Bedayat and Taleghani 2014; Khoshgoftar et al. 2007; Malekmotiei et al. 2013). Accordingly, the energy of an inclusion ΔF can be found by a simple integration scheme as (Mura 1987):

$$\Delta F = [\mathcal{E}(\nu) + \beta^2 \Psi(\nu)] \mu (\gamma^T)^2 \Omega, \quad (11)$$

in which γ^T and Ω are the predefined transformation shear strain and volume of an STZ, respectively, μ is the shear modulus of the sample, β is the dilatency parameter which can be found from the pressure sensitivity of the flow, $\Psi(\nu)$ is a function of the Poisson's ratio as $\Psi(\nu) = \frac{2(1+\nu)}{9(1-\nu)}$, and $\mathcal{E}(\nu)$ is a parameter which depends on both the Poisson's ratio and shape of the inclusion and varies between $\frac{7-5\nu}{30(1-\nu)}$ and 0.5 for spherical and flat ellipsoidal shapes, respectively. For the sake of simplicity, the STZs are assumed to be noninteracting inclusions (Argon 2013); hence, the total energy of the deformation can be obtained by superposition of the nucleation

energy of the individual STZs as inclusions (Bedayat and Taleghani 2014; Mura 1987; Shodja and Khorshidi 2013).

The use of the aforementioned form of energy for an inclusion which was proposed by Eshelby gives a constant value for the STZ's nucleation energy. However, as investigated by some researchers, this quantity evolves during the plastic deformation of amorphous solids (Argon 1979; Li et al. 2013). For a number of amorphous polymers, Oleinik (1991) obtained the total energy of deformation during their elastoplastic compression. According to his investigation, the total energy of deformation constantly increases with increasing the strain; however, the rate of the increment of the energy decreases beyond a certain strain (see Fig. 5 of Oleinik 1991). If one considers the aforementioned relation between the number of STZs and the plastic strain presented in Eq. 9, and the fact that the total plastic volume is equal to the total volume of the sample in uniaxial compression, the variation of the energy per STZ during the deformation can be obtained based on Oleinik's work, as presented in Fig. 2a for PMMA. Interestingly, the value of the energy of nucleation of a single STZ at yield which is found to be about 0.6 eV based on Oleinik's experiments is very close to that based on the authors' nanoindentation study on PMMA at room temperature (Malekmoitei et al. 2015). However, the latter method is not able to find the variation of this energy with increasing the strain.

Since the free volume void sites serve as the preferred spots for STZs formation, some investigators tried to link the variation of the STZs nucleation Helmholtz free energy to the evolution of the free volume during plastic strain to theoretically justify this phenomenon. The variation of the free volume holes' size and concentration with the strain were experimentally investigated by Hasan et al. (1993). Figure 2b shows the variation of the total free volume with the true uniaxial strain based on the work mentioned above and using Eqs. 5 and 6. Noticeably, the evolution of the free volume content in this figure is similar to that of the STZ nucleation energy as shown in Fig. 2a, but in the opposite direction. Both of these quantities change sharply upon yielding, and reach a stable value at a specific strain beyond which they show no significant change. This observation further supports the idea of existence of a relation between the free volume and shear transformation energy evolutions. Accordingly, some hypothetical models have been proposed to formulize this phenomenon. Notably for the case of MGs, Argon (1979) proposed a model which considered the dilatational component of the shear transformation to be stored as an excess free volume, the rate of the variation of which is calculated from creation-annihilation balance. In this way, the free volume creation takes place during the shear transformation events and its annihilation is mediated by the diffusive rearrangement of atoms. The free volume variation based on this model was recently used by Li et al. (2013) as a state variable for adjusting the STZ nucleation energy during the deformation. While the free volume content based on Argon's model alters with the dilatation component of the shear transformation, its effect needs to be considered on the dilatational part of the STZ nucleation energy as what was presented in Li et al. (2013). This, basically, means that the parameter β in Eq. 11 has to be considered as a function of free volume content. However, for the case of PGs, the parameter β is in the range of 0.2 (Malekmoitei et al. 2015;

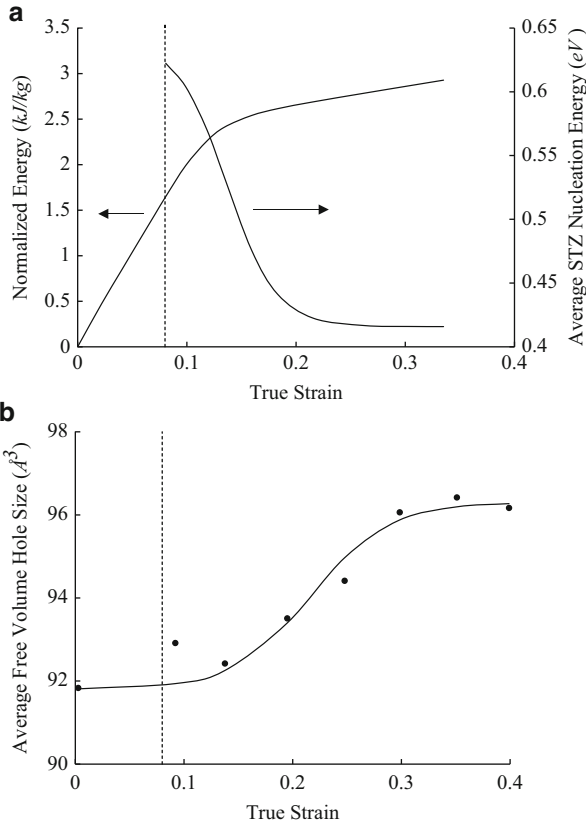


Fig. 2 (a) Variation of the stored internal energy in PMMA with true strain at room temperature for the compressive strain rate of 8.3×10^{-4} (s^{-1}) based on Oleinik (1991) (left axis) and the average nucleation energy per STZ calculated using Eq. 8 (right axis). (b) Data points show the variation of the average free volume size with true strain based on the compressive experiments of Hasan et al. (1993) on PMMA at room temperature and strain rate below β -transition, and the solid line is the \tanh interpolation as discussed later in this chapter. The dashed line in both figures is the projected yield strain (Reproduced from Voyiadjis and Samadi-Dooki (2016) with permission from AIP)

Samadi-Dooki et al. 2016) and its variation (which is squared in this equation) cannot bring about the variation of STZ nucleation energy required to capture the behavior presented in Fig. 2a. Hence, in the current study, the quantitative evolution of STZ nucleation energy with plastic strain is obtained directly from the analysis of Oleinik's experiments as shown in Fig. 2a.

Other parameters which can affect the nucleation energy of the STZ are the temperature and the thermal history of the PG sample. The nucleation Helmholtz free energy is directly dependent on the shear modulus of the material as understood from Eq. 11. Hence, it is reasonable that one assumes the alteration of the STZ

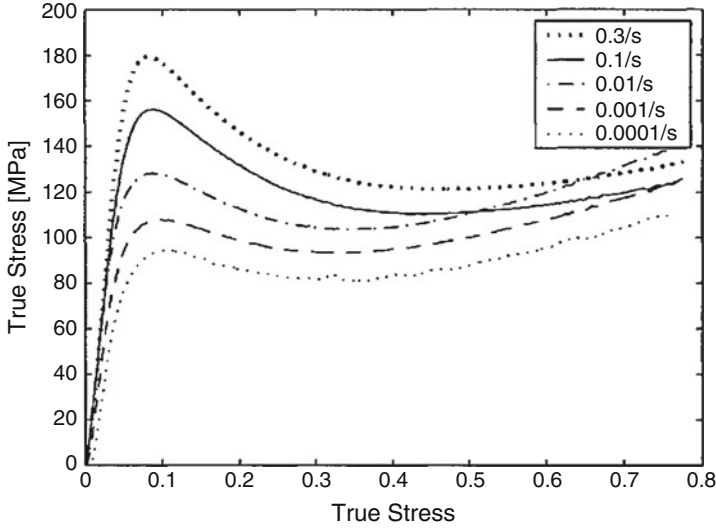


Fig. 3 Uniaxial compressive stress-strain curves of PMMA at different strain rates (Reproduced from Mulliken (2004) with permission from MIT)

nucleation energy to be a function of temperature in the same fashion of that for the shear modulus. This assumption is valid because other parameters in Eq. 11 can be assumed to be independent of temperature (see the discussion in Voyiadjis and Samadi-Dooki (2016) and the referencing to Holt (1968) and Roetling (1965)). The authors' previous investigation on PMMA shows that the STZ nucleation energy ΔF decreases with increasing temperature and almost vanishes at glass transition temperature T_g . This variation scheme is used in this study for investigating the temperature dependence of the flow. In addition, the authors' previous studies demonstrate a slight increase of the STZ nucleation energy with a short annealing time. Since the free volume content of the PG decreases with thermal aging, and an inverse relation between the free volume and shear transformation energy evolution seems to exist, the increase of the STZ nucleation energy could be qualitatively justified. The quantitative investigation of this phenomenon, however, is deferred to a forthcoming study by the authors.

Concluding from these remarks, the STZ nucleation energy in a PG can be taken as a function of temperature T , shear plastic strain γ^p , and thermal history of the sample θ as:

$$\Delta F = \Delta F(\gamma^p, T, \theta). \quad (12)$$

Finally, in light of Eqs. 10 and 12, and assuming $\gamma^T \Omega \tau^p \gg 2k_B T$ for PGs at temperatures below their T_g (Argon 2013), Eq. 1 can be reorganized at a constant strain rate as:

$$\tau^p(\gamma^p, \dot{\gamma}^p, T, \theta) = \frac{2k_B T}{\gamma^T \Omega} \left[\ln \left(\frac{2\dot{\gamma}^p}{\dot{\gamma}_0^p} \right) + \frac{\Delta F(\gamma^p, T, \theta)}{k_B T} + \frac{k\gamma^p}{f(\gamma^p, T, \theta)} \right]. \quad (13)$$

Equation 13 represents the explicit plastic shear stress-strain relation of a PG. To compare the results based on the current theory and those based on experiments, the shear stress and strain quantities can be simply converted to uniaxial parameters according to the von Mises' equivalency concept as:

$$\sigma^p = \sqrt{3}\tau^p \quad (14a)$$

and

$$\gamma^p = \sqrt{3}\epsilon^p, \quad (14b)$$

with σ^p and ϵ^p representing the uniaxial plastic stress and strain, respectively.

To complete the model, the viscoelastic preyield behavior is considered as a simple Maxwell viscoelastic material, the governing differential equation of which reads as:

$$\dot{\epsilon}^{ve}(t) = \frac{\dot{\sigma}^{ve}(t)}{E} + \frac{\sigma^{ve}(t)}{\eta}, \quad (15)$$

in which $\sigma^{ve}(t)$ and $\epsilon^{ve}(t)$ are the preyield viscoelastic stress and strain, respectively, $\dot{\epsilon}^{ve}(t)$ is the strain rate, E is the Young's modulus which is temperature dependent, η is the viscosity which is strain rate and temperature dependent (Kobayashi et al. 2001; Sun 2007), and t is time. The solution of Eq. 15 for constant strain rates, where $\dot{\epsilon}^{ve}(t) = \dot{\epsilon}^{ve}$, is readily found as:

$$\sigma^{ve}(t) = \dot{\epsilon}^{ve}\eta + A_1 e^{-\frac{Et}{\eta}}, \quad (16)$$

where A_1 is the constant of the solution which is easily obtained from the boundary condition of $\sigma^{ve}(0) = 0$. Assuming that the contribution of the elastic component of the strain rate during the postyield deformation is negligible (Hasan et al. 1993), Eqs. 16 and 13 can be considered to independently represent the stress-strain behavior in the preyield and postyield regions.

Results and Discussion

To evaluate the accuracy of the current model, the large deformation of PMMA as a glassy polymer is investigated using the proposed formalism which is compared with the available experimental data for this material. PMMA is a transparent polymer which is of both industrial and medical importance. Accordingly, it has attracted many researchers to study its manufacturing process and characteristic properties, including its mechanical behavior. Notably, Ghadipasha et al. (2016) developed a framework for investigating the polymerization of this polymer using

integrated simulation, optimization, and online feedback control; Gunel and Basaran (2009, 2010, 2011a, b, 2013) developed comprehensive experimental, analytical, and numerical approaches for scrutinizing the deformation of PMMA and its composites with taking into account the effect of temperature and damage evolution; and Nasraoui et al. (2012) studied its mechanical behavior dependence on the strain rate and temperature both experimentally and analytically.

In this study, effects of strain rate, temperature, and thermal history of the sample on the postyield stress-strain behavior of PMMA are obtained based on the proposed model and compared with the experimental observations. The input parameters of the model are adopted from relevant experimental studies. In particular, the numerical values for the characteristic parameters of the STZ in PMMA, like its strain, volume, nucleation energy at yield, and the variation of this energy with temperature and thermal history of the sample, are extracted from the experimental evaluation of Malekmtotiei et al. (2015). In addition, the evolution of the STZ nucleation energy during inelastic deformation is taken into account with considering the calculated values based on the measurements by Oleinik (1991) as presented in Figs. 1 and 2a. The variation of the free volume content of PMMA with strain, temperature, and thermal history of the sample is also calculated using the results of the work by Hasan et al. (1993) and Hristov et al. (1996).

Effect of the Rate

When tested at a wide range of strain rates, the variation of the yield stress of glassy polymers shows two distinct regions, each of which is almost linear with respect to the logarithm of the strain rate. The separation point might be called the β -transition point according to the study by Mulliken and Boyce (2006), and the trend of the variation of yield point is sharper in post- β -transition area compared to the pre- β -transition region. While the β -transition strain rate for PMMA in uniaxial loading condition was shown to be temperature dependent (Roetling 1965), it is generally considered to be about 0.01 s^{-1} at room temperature (Malekmtotiei et al. 2015; Mulliken and Boyce 2006). The distinct pre- and post- β -transition behaviors are not just limited to the yield point of PMMA. The overall plastic deformation of this polymer shows rate dependence (as seen in Fig. 3) such that the postyield behavior might be divided into two distinguishable parts. Obviously, in each region, only the yield stress is strain rate sensitive, and the curves for the postyield behavior are parallel. In addition, while the steady-state is reached at the true strain of about 0.3 for pre- β -transition strain rates, it is shifted to the strain of about 0.45 for the strain rates of above the β -transition. In this study, we limit the investigation of the effect of the rate on the PMMA postyield behavior to the sub- β -transition strain rates. The reason for this adjustment is that the numerical values pertaining to the microstructural properties of the STZs, free volume evolution, and stored deformation energy variation of PMMA are given or calculated for the strain rates below the β -transition (in Malekmtotiei et al. 2015; Hasan et al. 1993; Oleinik 1991, respectively).

During the plastic deformation at a constant strain rate, the free volume content of the PG increases as previously shown in Fig. 2b. The trend of the increment is such that the variation is insignificant prior to the yield, then it increases sharply until it gets to a stabilized value; this variation is acceptably interpolated by hyperbolic tangent (*tanh*) function. Interestingly, this function is compatible with the solution of the differential equation for the variation of c_f with the plastic strain at a constant strain rate which is presented in De Hey et al. (1998) as:

$$\frac{dc_f}{d\gamma} = -k_r c_f (c_f - c_{f,eq}) + a_f, \quad (17)$$

with k_r standing for a rate constant which is also dependent on the activation energy, a_f for a temperature-dependent parameter, and $c_{f,eq}$ for the equilibrium value of the fraction of the potential jump sites. In addition, since the evolution of the STZ nucleation energy with the plastic strain follows the same trend as that for the free volume (see Fig. 2), the same *tanh* interpolation format with required adjustments is used to represent its variation.

In Table 1 the numerical values for model parameters pertaining to PMMA at three sub- β -transition strain rates and at room temperature are shown. Accordingly, the true stress versus true strain curves based on the proposed model are obtained and compared to the numerical values based on the experiments of Mulliken (2004) as shown in Fig. 4. A very important fact which is indicated in this figure is that all curves are almost parallel in the postyield region, and the yield stress is only affected by the strain rate. Luckily, the structure of Eq. 13 readily reveals the reason. In this equation ΔF and f are independent of the strain rate, and the only factor which varies with this parameter is the first term in the right-hand side; while this term is independent of the plastic strain, it causes a constant jump in the curves proportional to the logarithm of the strain rate. In addition, Fig. 4 demonstrates that the proposed model accurately captures the primary softening and secondary hardening of PMMA. When the results are scrutinized in detail, it is observed that the secondary hardening is close to a linear variation. Since the free volume and the STZ nucleation energy converge to their respective plateaus beyond the steady state

Table 1 Numerical values for model parameters pertaining to PMMA. In this table, η is the viscosity, E is the Young's modulus, ΔF_0 is the STZ nucleation energy at yield, and $\gamma^T \Omega$ is the STZ size scale

Parameter Strain rate (s ⁻¹)	η (Pa.s)	E (GPa)	ΔF_0 (eV)	$\gamma^T \Omega$ (nm ³)
0.0001	10 ¹²	4.8	0.62	3.7
0.001	10 ¹¹			
0.01	1.2 × 10 ¹⁰			
Reference(s)	Sun (2007) and Kobayashi et al. (2001)	Gall and McCrum (1961)	Malekmtotiei et al. (2015)	Malekmtotiei et al. (2015)

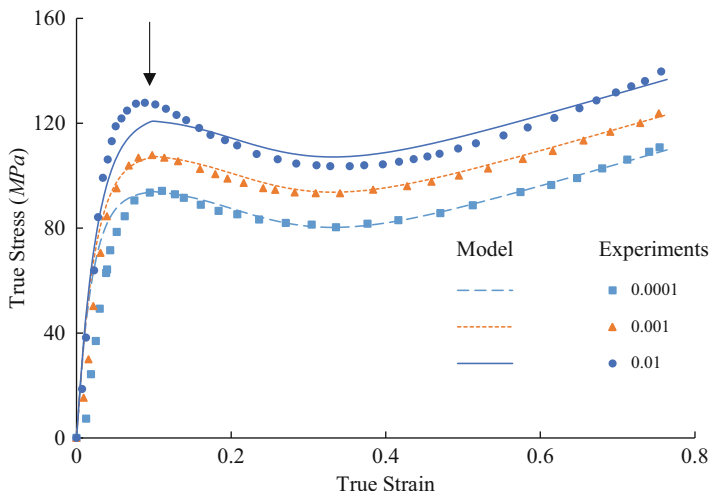


Fig. 4 True stress versus true strain curves of PMMA deformation at three different strain rates based on the proposed model as compared to the experiments of Mulliken (2004). The vertical arrow shows the yield point (Reproduced from Voyiadjis and Samadi-Dooki (2016) with permission from AIP)

strain, according to the last term of Eq. 13 the stress should increase with the number of STZs which has a linear variation with strain. While this behavior is very close to the experimental observations, some studies suggest that at higher plastic strains, PMMA might show a higher order hardening rather than a linear one (Arruda et al. 1995). This latter could be a result of chain alignment inside or on the border of the STZs which can increase the overall material resistance to further deformation. As another significant observation, at the strain rate of 0.01 s^{-1} which is close to the β -transition strain rate, the behavior of the material starts to deviate from that at sub- β -transition strain rates and also the viscoelastic-viscoplastic formulation developed in the study. The proper treatment of this phenomenon requires the consideration of the adiabatic heat generation in the PG at higher deformation rates, and is deferred to a forthcoming study by the authors.

Effect of the Temperature

The thermoplastics response to mechanical loading is undeniably affected by the temperature at which it is loaded. The formalism presented in this study considers the effect of the temperature on pre- and postyield behavior of PGs as reflected in Eqs. 13 and 16, respectively. For the preyield response, the temperature is a determining factor such that both viscosity and elastic modulus in Eq. 16 are temperature dependent, and hence temperature variation alters the viscoelastic response of the polymer. On the other hand, in the constitutive relation of Eq. 13

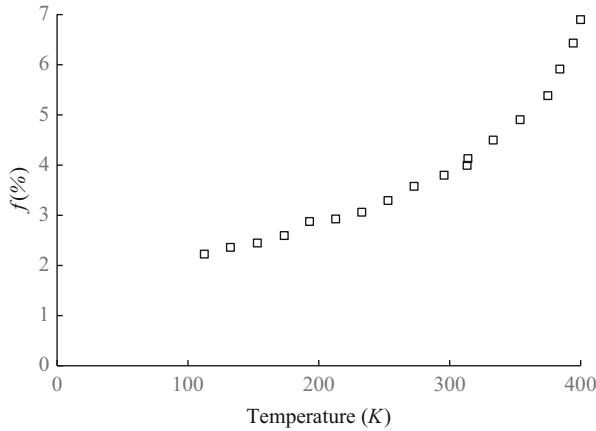


Fig. 5 Variation of the fraction of the free volume of PMMA with temperature (Reproduced from Hristov et al. (1996) with permission from ACS)

Table 2 Numerical values for temperature-dependent model parameters. The free volume content at yield is f_0 . Other parameters hold the same definition as presented in the caption of Table 1

Parameter Temp. (K)	η (Pa.s)	E (GPa)	ΔF_0 (eV)	f_0 (%)
300	10^{11}	4.8	0.62	3.8
325	7×10^{10}	3.6	0.47	4.5
375	3.5×10^{10}	1.7	0.33	5.4
Reference(s)	Sun (2007) and Kobayashi et al. (2001)	Gall and McCrum (1961)	Malekmtotiei et al. (2015)	Hristov et al. (1996)

for the postyield stage, the STZ nucleation energy ΔF and free volume fraction f are both functions of temperature (see Eqs. 12 and 7, respectively). Since ΔF is a function of shear modulus μ , it can be assumed to vary with temperature in the same fashion as the variation of shear modulus with temperature. In addition, since both the free volume hole size and concentration have been shown to vary almost linearly with temperature (Hasan et al. 1993), the free volume fraction f is expected to vary with a rate faster than a linear variation with temperature according to Eq. 6. This higher order variation was beheld by Hristov et al. (1996) for a number of glassy polymers including PMMA as shown in Fig. 5.

The numerical values of the temperature-dependent parameters involved in Eqs. 13 and 16 are presented in Table 2. Using these values, the true stress versus true strain curves of PMMA at the compressive strain rate of 0.001 s^{-1} are obtained for three different temperatures as depicted in Fig. 6 which shows a good correlation with the experimental observations of Arruda et al. (1995). The results of Fig. 6 demonstrate two notable alterations of the PGs' behavior with variation of the temperature. Firstly, with increasing the temperature the yield stress decreases; this

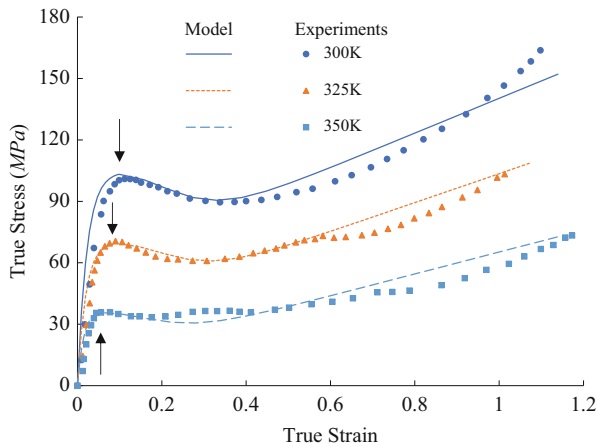


Fig. 6 True stress versus true strain curves of PMMA at the compressive strain rate of 0.001 s^{-1} for three different temperatures based on the proposed model and experiments of Arruda et al. (1995). Both yield strain and stress reduce with increasing the temperature (Reproduced from Voyiadjis and Samadi-Dooki (2016) with permission from AIP)

is a direct result of the STZs nucleation energy reduction at higher temperature as extensively discussed in Malekmoitei et al. (2015). Secondly, the variation of the temperature also changes the postyield behavior such that higher temperature PMMA undergoes primary softening and secondary hardening which are dampened compared to lower temperature behavior; this phenomenon is also well captured by the proposed model. As a justification, it could be mentioned that the higher order variation of f with temperature in the denominator of the third term in the right-hand side of Eq. 13 is faster than the linear variation of the numerator. As a result, the variation of the stress with plastic strain is damped by the effect of the increased temperature.

Effect of the Thermal History

Heat treatment is a process for enhancing the mechanical resistance of many solid state materials by relaxing the microstructural disturbances which might exist in these solids as residual stress or strain, or extra cavities and defects. For glassy polymers, such extra defects might exist if the cooling rate from molten state to solid state is relatively high (quenching). Since instabilities which are the result of unrelaxed microstructure can negatively affect the mechanical properties of the PGs, they could be physically aged at temperatures close to their glassy temperature for several hours in order to regain their nominal strength. This process is called annealing at it is shown to increase the yield stress of PGs (Chen and Schweizer 2011; Hutchinson et al. 1999; Nanzai et al. 1999). Nevertheless, both quenched

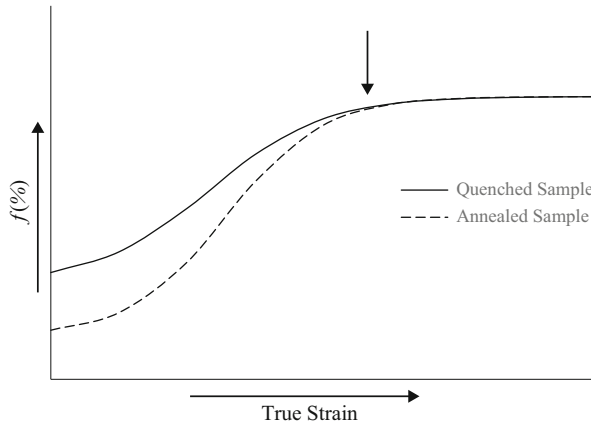


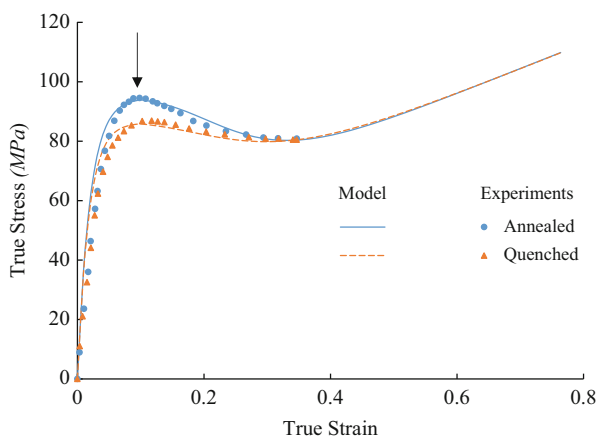
Fig. 7 Schematic representation of the variation of the free volume fraction with strain for annealed and quenched samples based on Hasan et al. (1993). Vertical arrow shows the steady-state strain (Reproduced from Voyiadjis and Samadi-Dooki (2016) with permission from AIP)

and annealed samples of the same polymer are shown to reach the same postyield steady-state stress which takes place at the same strain, and afterwards they both exhibit identical secondary hardening (Hasan et al. 1993; Klompen et al. 2005; Xiao and Nguyen 2015). This phenomenon seems to be correlated with the free volume evolution of the polymeric samples with different thermal histories. For example, the investigation by Hasan et al. (1993) shows that while the annealed PMMA sample possesses a lower free volume content compared to the quenched one, the free volume fraction for both of the samples reaches the same value when they are deformed up to the steady-state plastic strain. This is schematically shown in Fig. 7. On the other hand, since it is believed that there exists an inverse correlation between the free volume content and the STZ nucleation energy, one might expect a higher STZ activation energy at yield for annealed sample compared to the quenched one. The experimental evaluation of the STZ nucleation energy at yield for PMMA and Polycarbonate (PC) with different thermal histories confirms this hypothesis with revealing a slightly bigger STZ nucleation energy for annealed samples (Malekmoitei et al. 2015; Samadi-Dooki et al. 2016).

According to the aforementioned correlation between the free volume content and the STZ nucleation energy, it seems to be reasonable to assume that the STZ nucleation energy of quenched and annealed PMMA samples also reaches the same value at the steady-state strain and remains identical after that. Hence, with considering the STZ activation energy variation with the strain as presented in Fig. 2a, the *tanh* interpolation function for this quantity could be utilized with the required adjustments for quenched and annealed samples. Accordingly, and with using the numerical values for model parameters as presented in Table 3, the true stress versus true strain curves for samples with different thermal histories could be obtained as presented in Fig. 8. The prediction of the model for considering

Table 3 Numerical values for model parameters for PMMA with different thermal histories (the value of η for quenched sample is assumed for curve fitting purpose)

Parameter Thermal History	η (Pa.s)	E (GPa)	ΔF_0 (eV)	$\gamma^T \Omega$ (nm ³)	f_0 (%)
Annealed	10^{12}	4.8	0.62	3.7	3.8
Quenched	0.86×10^{12}		0.58	3.6	3.9
Reference(s)	Sun (2007), Kobayashi et al. (2001)	Gall and McCrum (1961)	Malekmoitei et al. (2015)	Malekmoitei et al. (2015)	Hristov et al. (1996) and Hasan et al. (1993)

**Fig. 8** True stress versus true strain of PMMA for annealed and quenched samples based on the proposed model and experiments of Hasan and Boyce (1995) at strain rate of 0.0001 s^{-1} . The vertical arrow shows the yield (Reproduced from Voyiadjis and Samadi-Dooki (2016) with permission from AIP)

the thermal history of PMMA on its stress-strain behavior is acceptably close to the experimental observation by Hasan and Boyce (1995) for the softening region. The secondary hardening is predicted to be identical for the samples with different thermal histories which is in agreement with the experimental observation for other glassy polymers like PC (see Fig. 2 of Hasan et al. 1993).

Implication of the Model for Shear Banding and Indentation Size Effect in PGs

Since the model parameters in this study are obtained from compressive experiments, the authors were careful to make the comparisons with compressive behavior of PMMA where no localization effect is expected to exist. However,

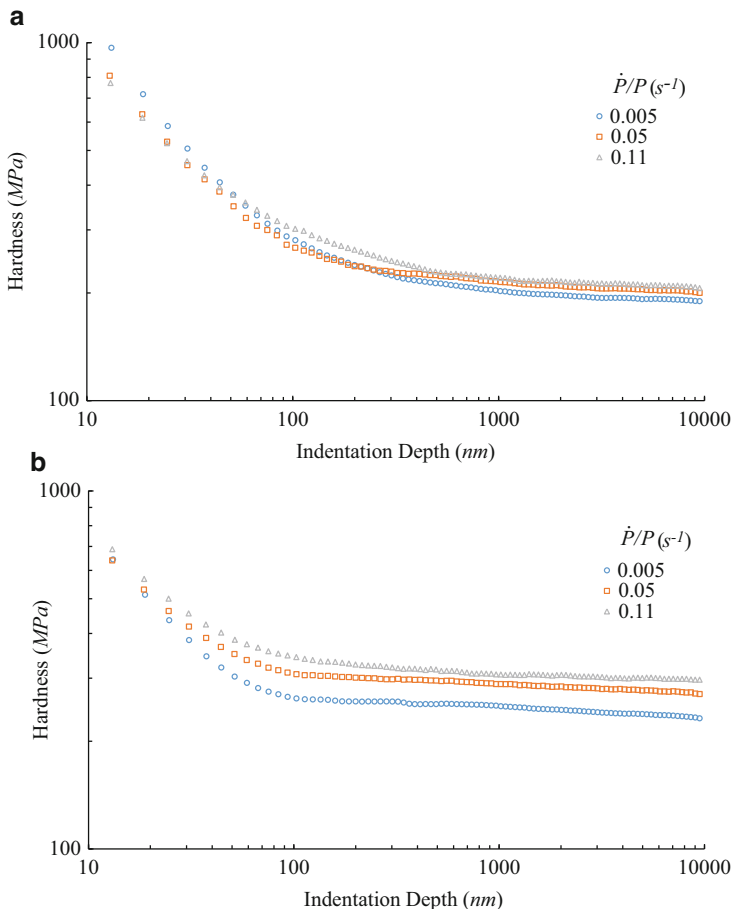


Fig. 9 Variation of the hardness with indentation depth at three different rates for (a) PC and (b) PMMA (Reproduced from Voyiadjis and Malekmotiei (2016) with permission from Wiley Online Library)

the localization can happen in polymeric samples under more complex loading conditions with formation of shear bands (Anand and Spitzig 1982; Li and Buckley 2009; Voyiadjis et al. 2014) or where the strain gradient can generate an additional hardening which, for example, emerges as indentation size effect (ISE) in nanoindentation of glassy polymers (Malekmotiei et al. 2015; Samadi-Dooki et al. 2016; Voyiadjis and Malekmotiei 2016; Voyiadjis et al. 2014). The latter is the increased hardness at shallow indentation in nanoindentation experiments with self-similar tips like the Berkovich tip (see Fig. 9). Since the proposed model takes into account the free volume and STZs evolutions, it can be used to evaluate these phenomena. For example, the formation of shear bands in PGs can be thought of as coalescence of STZ within a small region where extra defects (free volume) exist. In addition, the ISE in polymers might also be formulated with considering

the probability of presence of a free volume site in the small anticipated plastic region which can accommodate the required number of STZs to bring about the plastic deformation. While modeling these effects is an ongoing topic of study by the authors, they differ further discussions to their forthcoming publications.

Concluding Remarks

In this study, the plastic deformation of glassy polymers is formulated based on the evolutions of the free volume and the shear transformation nucleation energy during the course of large inelastic deformation. In order to be used as input parameters, these variations were quantified based on previous experimental observations. The reason for this strategy is that the authors have found that the direct use of the formulations developed for the evolution of these two parameters in MGs cannot account for their variations in polymeric glasses. To test the accuracy of the model, effects of the strain rate, temperature, and thermal history of the samples on the postyield behavior of PMMA, as a widely used PG, have been analyzed and compared with the previous experimental observations. The results show acceptable consonance with experiments for both the primary postyield softening and the secondary hardening behavior of PMMA. Although the authors have made some simplifying assumptions, the proposed model is promising in a sense that it captures the softening-hardening deformation behavior of PGs with a single formula. This suggests that the overall plastic deformation behavior of PGs is a unified mechanism. This hypothesis, which is well proved in this study, is fundamentally different from previous studies which suggest different mechanisms for the postyield softening-hardening behavior of PGs. With additional developments, the proposed model can stand as a conclusive theory for plastic deformation of amorphous polymers.

References

- L. Anand, W. Spitzig, *Acta Metall.* **30**, 553 (1982)
- A. Argon, *Philos. Mag.* **28**, 839 (1973)
- A. Argon, *Acta Metall.* **27**, 47 (1979)
- A.S. Argon, *Strengthening Mechanisms in Crystal Plasticity* (Oxford University Press, Oxford, 2008)
- A.S. Argon, *The Physics of Deformation and Fracture of Polymers* (Cambridge University Press, New York, 2013)
- A. Argon, M. Demkowicz, *Philos. Mag.* **86**, 4153 (2006)
- A. Argon, M. Demkowicz, *Metall. Mater. Trans. A* **39**, 1762 (2008)
- E.M. Arruda, M.C. Boyce, in *Anisotropy and Localization of Plastic Deformation*, ed. by J-P. Boehler, A.S. Khan (Elsevier Applied Science, London and New York, 1991), p. 483
- E.M. Arruda, M.C. Boyce, H. Quintus-Bosz, *Int. J. Plast.* **9**, 783 (1993)
- E.M. Arruda, M.C. Boyce, R. Jayachandran, *Mech. Mater.* **19**, 193 (1995)
- H. Bedayat, A.D. Taleghani, *Mech. Mater.* **69**, 204 (2014)
- M.C. Boyce, D.M. Parks, A.S. Argon, *Mech. Mater.* **7**, 15 (1988)

- K. Chen, K.S. Schweizer, *Macromolecules* **44**, 3988 (2011)
- M.H. Cohen, D. Turnbull, *J. Chem. Phys.* **31**, 1164 (1959)
- P. De Hey, J. Sietsma, A. Van Den Beukel, *Acta Mater.* **46**, 5873 (1998)
- J.D. Eshelby, *Proc. Roy. Soc. London Ser. A* **241**, 376 (1957)
- H. Eyring, *J. Chem. Phys.* **4**, 283 (1936)
- M.L. Falk, *Science* **318**, 1880 (2007)
- M. Falk, J. Langer, L. Pechenik, *Phys. Rev. E* **70**, 011507 (2004)
- K. Flores, D. Suh, R. Dauskardt, P. Asoka-Kumar, P. Sterne, R. Howell, *J. Mater. Res.* **17**, 1153 (2002)
- W. Gall, N. McCrum, *J. Polym. Sci.* **50**, 489 (1961)
- N. Ghadipasha, A. Geraili, J.A. Romagnoli, C.A. Castor, M.F. Drenski, W.F. Reed, *Processes* **4**, 5 (2016)
- E. Gunel, C. Basaran, *Mater. Sci. Eng. A* **523**, 160 (2009)
- E. Gunel, C. Basaran, *J. Eng. Mater. Technol.* **132**, 031002 (2010)
- E. Gunel, C. Basaran, *Mech. Mater.* **43**, 979 (2011a)
- E. Gunel, C. Basaran, *Mech. Mater.* **43**, 992 (2011b)
- E. Gunel, C. Basaran, *Mech. Mater.* **57**, 134 (2013)
- O. Hasan, M. Boyce, *Polym. Eng. Sci.* **35**, 331 (1995)
- O. Hasan, M. Boyce, X. Li, S. Berko, *J. Polym. Sci. B Polym. Phys.* **31**, 185 (1993)
- P. de Hey, J. Sietsma, A. Van Den Beukel, *Mater. Sci. Eng. A* **226**, 336 (1997)
- J. Ho, L. Govaert, M. Utz, *Macromolecules* **36**, 7398 (2003)
- D. Hofmann, M. Heuchel, Y. Yampolskii, V. Khotimskii, V. Shantarovich, *Macromolecules* **35**, 2129 (2002)
- D.L. Holt, *J. Appl. Polym. Sci.* **12**, 1653 (1968)
- H. Hristov, B. Bolan, A. Yee, L. Xie, D. Gidley, *Macromolecules* **29**, 8507 (1996)
- J. Hutchinson, S. Smith, B. Horne, G. Gourlay, *Macromolecules* **32**, 5046 (1999)
- Y. Jean, *Microchem. J.* **42**, 72 (1990)
- Y. Jean, J.D. Van Horn, W.-S. Hung, K.-R. Lee, *Macromolecules* **46**, 7133 (2013)
- M. Khoshgoftar, S. Najarian, F. Farmanzad, B. Vahidi, F. Ghomshe, *Am. J. Appl. Sci.* **4**, 918 (2007)
- E. Klompen, T. Engels, L. Govaert, H. Meijer, *Macromolecules* **38**, 6997 (2005)
- H. Kobayashi, H. Takahashi, Y. Hiki, *J. Non-Cryst. Solids* **290**, 32 (2001)
- D.C. Lam, A.C. Chong, *J. Mater. Res.* **14**, 3784 (1999)
- J. Langer, L. Pechenik, *Phys. Rev. E* **68**, 061507 (2003)
- H. Li, C. Buckley, *Int. J. Solids Struct.* **46**, 1607 (2009)
- L. Li, E. Homer, C. Schuh, *Acta Mater.* **61**, 3347 (2013)
- L. Maleknotiei, F. Farahmand, H.M. Shodja, A. Samadi-Dooki, *J. Biomech. Eng.* **135**, 041004 (2013)
- L. Maleknotiei, A. Samadi-Dooki, G.Z. Voyiadjis, *Macromolecules* **48**, 5348 (2015)
- A.D. Mulliken, *Low to high strain rate deformation of amorphous polymers: experiments and modeling* (Massachusetts Institute of Technology, Massachusetts, 2004)
- A. Mulliken, M. Boyce, *Int. J. Solids Struct.* **43**, 1331 (2006)
- T. Mura, in *Micromechanics of Defects in Solids* (Martinus Nijhoff Publishers, The Netherlands, 1987)
- Y. Nanzai, A. Miwa, S.Z. Cui, *JSME Int. J. Ser. A Solid Mech. Mater. Eng.* **42**, 479 (1999)
- M. Nasraoui, P. Forquin, L. Siad, A. Rusinek, *Mater. Des.* **37**, 500 (2012)
- E. Oleinik, in *High Performance Polymers*, ed. by E. Baer, A. Moet (Hanser, New York, 1991), p. 79
- E. Oleinik, S. Rudnev, O. Salamatina, *Polym. Sci. Ser. A* **49**, 1302 (2007)
- S. Pauly, S. Gorantla, G. Wang, U. Kühn, J. Eckert, *Nat. Mater.* **9**, 473 (2010)
- T. Ree, H. Eyring, *J. Appl. Phys.* **26**, 793 (1955)
- R.E. Robertson, *J. Chem. Phys.* **44**, 3950 (1966)
- J. Roetting, *Polymer* **6**, 311 (1965)
- A. Samadi-Dooki, L. Maleknotiei, G.Z. Voyiadjis, *Polymer* **82**, 238 (2016)
- H. Shodja, A. Khorshidi, *J. Mech. Phys. Solids* **61**, 1124 (2013)

- H. Shodja, M. Tabatabaei, A. Ostadhossein, L. Pahlevani, *Open Eng.* **3**, 707 (2013)
- F. Spaepen, *Acta Metall.* **25**, 407 (1977)
- Y. Sun, *J. Cent. S. Univ. Technol.* **14**, 342 (2007)
- L.A. Utracki, A.M. Jamieson, *Polymer Physics: From Suspensions to Nanocomposites and Beyond* (Wiley, New York, 2011)
- G.Z. Voyiadjis, L. Malekmoitei, *J. Polym. Sci. Part B: Polym. Phys.* **54**, 2179 (2016)
- G.Z. Voyiadjis, A. Samadi-Dooki, *J. Appl. Phys.* **119**, 225104 (2016)
- G.Z. Voyiadjis, M. Yaghoobi, *Mater. Sci. Eng. A* **634**, 20 (2015)
- G.Z. Voyiadjis, A. Shojaei, N. Mozaffari, *Polymer* **55**, 4182 (2014)
- Y. Wang, H. Nakanishi, Y. Jean, T. Sandreczki, *J. Polym. Sci. B Polym. Phys.* **28**, 1431 (1990)
- D. Wang, Z. Zhu, R. Xue, D. Ding, H. Bai, W. Wang, *J. Appl. Phys.* **114**, 173505 (2013)
- R. Xiao, T.D. Nguyen, *J. Mech. Phys. Solids* **82**, 62 (2015)
- A.R. Yavari, A. Le Moulec, A. Inoue, N. Nishiyama, N. Lupu, E. Matsubara, W.J. Botta, G. Vaughan, M. Di Michiel, Å. Kvick, *Acta Mater.* **53**, 1611 (2005)



Baoxing Xu, Xi Chen, and Zhufeng Yue

Contents

Introduction	402
Mechanics Theory of Indentation Fatigue	404
Indentation Load-Depth Curve	404
Indentation Stress Intensity Factor	406
Indentation Fatigue Deformation	409
Indentation Fatigue Depth Propagation Law	409
Computational Validation	410
Experimental Validation	413
Extend to Overloading and Underloading	416
Microstructural Observation	419
Discussion on Deformation Mechanism	421
Indentation Fatigue Damage	422
Indentation Fatigue Strength Law	422
Experimental Validation	423
Concluding Remarks	427
References	427

B. Xu (✉)

Department of Mechanical and Aerospace Engineering, University of Virginia, Charlottesville, VA, USA

e-mail: bx4c@virginia.edu

X. Chen

Department of Earth and Environmental Engineering, Columbia Nanomechanics Research Center, Columbia University, New York, NY, USA

e-mail: xichen@columbia.edu

Z. Yue

Department of Engineering Mechanics, Northwestern Polytechnical University, Xi'an, Shaanxi, China

e-mail: zfyue@nwpu.edu.cn

Abstract

Instrumented indentation has been widely used in the determination of mechanical properties of materials due to its fast, simple, precise, and nondestructive merits over the past few years. In this chapter, we will present an emerging indentation technique, referred to as indentation fatigue, where a fatigue load is applied on a sample via a flat punch indenter, and establish the framework of mechanics of indentation fatigue to extract fatigue properties of materials. Through extensive experimental, theoretical, and computational investigations, we demonstrate a similarity between the indentation fatigue depth propagation and the fatigue crack growth, and propose an indentation fatigue depth propagation law and indentation fatigue strength law to describe indentation fatigue-induced deformation and failure of materials, respectively. This study provides an alternative approach for determining fatigue properties, as well as for studying the fatigue mechanisms of materials, especially for materials that are not available or feasible for conventional fatigue tests.

Keywords

Indentation · Fatigue loading · Indentation depth propagation · Crack propagation · Strength

Introduction

Fatigue is a process of accumulations of material deformation and damage due to a repeated loading and unloading and is considered as one of major threats to the mechanical integrity of materials and structures. As a technical problem, the study of fatigue can be traced to as early as the mid-nineteenth century with the data collection of failure of railroad system by Germany technologist A Wohler (Suresh 1998). Since then, it has attracted tremendous attention over the last century, and great achievements have been made from macroscaled phenomena to micro-/nanoscaled deformation mechanisms (Estrin and Vinogradov 2010; Connolley et al. 2005), from qualitative descriptions to quantitative predictions (Rao and Farris 2008; Fleck and Smith 1984), and from engineering metals to biological and soft materials (Teoh 2000; Dirks et al. 2013; Tang et al.). Meanwhile, with the ever-growing applications of small structures, including nano-/micro-electromechanical systems (N/MEMS) and thin films over the last decades and emerging low-dimensional nanomaterials such as nanowires, nanofibers, and graphene, their properties usually exhibit differences than their materials counterpart at the macroscale (Alsem et al. 2007, 2008; Höppel et al. 2009; Li et al. 2003, 2014; Luo et al. 2015; Lee et al. 2008; Zhu et al. 2007), and the characterization of their mechanical properties in particular fatigue properties is challenging the traditional uniaxial mechanical testing technique that is mainly based on standard “dog-bone” shaped specimens. The development of an alternative testing technique that could help to address these challenges is highly desired.

Indentation, which originates from the hardness measurement, provides a compelling solution to measure mechanical properties of materials down to the nanosize (Cheng and Cheng 2004; Fischer-Cripps 2000). The merits of indentation include minor requirements of testing materials, easy to conduct, and high resolutions of measurement. During the experiment, the indentation load and displacement can be recorded continuously with the penetration of the indenter to samples, and even deformation of materials can also be observed when the testing platform is integrated with advanced observational facilities such as transmission electron microscopy (TEM) and scanning electron microscope (SEM), usually referred to as *in situ* TEM-indentation (Warren et al. 2007) and SEM-indentation (Nowak et al. 2010), respectively. Significant efforts have been made to extract mechanical properties of materials from indentation data including elastic properties (Oliver and Pharr 2004; Yang et al. 2016; Li and Bhushan 2002; Dao et al. 2001; Lan and Venkatesh 2007; Jiang et al. 2009), plastic properties (Cheng and Cheng 1998; Chen et al. 2007; Xu and Chen 2010; Lee et al. 2010; Bucaille et al. 2003), fracture properties (Lawn and Wilshaw 1975; Tang et al. 2008; Xia et al. 2004; Sakaguchi et al. 1992; Miranzo and Moya 1984; Quinn and Bradt 2007), creep properties (Xu et al. 2008a; Cheng and Cheng 2001; Li et al. 1991; Yang and Li 1995; Chen et al. 2010; Stone et al. 2010), and relaxation properties (Xu et al. 2008b; Baoxing et al. 2010; Chu and Li 1977, 1980a; Hu et al. 2011; Chan et al. 2012) with a broad range of materials from hard metals, to composites, to biological materials, to soft matters. Usually, these properties can be determined from the measured indentation load and displacement curves and deformation profiles of indented materials at monotonic loading conditions. However, fatigue properties such as fatigue strength and fatigue-induced crack growth that rely on fatigue/cyclic loading conditions usually cannot be measured.

Indentation that is conducted under a cyclic loading with cyclic numbers >1000 is referred to indentation fatigue here with an emphasis on measurement of fatigue properties of materials. An early study by Li and Chu found that the flat punch indenter could continue to penetrate into the β -tin single crystal with applied cycles, and the recorded curves of indentation displacement-cyclic numbers were similar to that of fatigue crack propagation with the absence of third stage (Li and Chu 1979). With the same β -tin single crystal, later on, Chu and Li changed the fatigue conditions to fatigue spectrum with interruption by a peak loading and further found the similarity with that of fatigue crack propagation with a delayed retardation of indentation depth propagation caused by a peak loading (Chu and Li 1980b). By employing a Vickers diamond indenter, Kaszynski et al. investigated the indentation fatigue response of 316 L stainless steel and found that the indentation depth showed an approximately linear increase with the logarithm of the number of cycles (Kaszynski et al. 1998). This difference is believed to be led by the geometric shape of indenters and associated change of contact areas between indenter and materials in experiments (Li 2002; Yang and Li 2013). The similarities between indentation fatigue and fatigue crack propagation provide a critical hint to probe the fatigue properties of materials through instrumented indentation technique

(Yang and Li 2013). However, why do such similarities exist? What is the deformation mechanism of the evolution of indentation depth propagation under a fatigue loading? What fatigue properties can be determined and how? How to quantitatively describe the depth propagation of indentation fatigue? What is the quantitative relationship between the depth propagation of indentation fatigue and fatigue crack propagation of conventional fatigue test? Over the last 10 years, we have been actively working on the study of indentation fatigue by integrating tools of experiments, theories, and computations to address these questions (Xu et al. 2010, 2007a, b, 2009; Xu and Yue 2006, 2007).

In this book chapter, we will summarize several important results based on our progresses and present quantitative mechanics descriptions and validations of indentation fatigue from indentation fatigue depth propagation to indentation fatigue strength of materials. In section “[Mechanics Theory of Indentation Fatigue](#),” an elastic indentation fatigue mechanics model will be first developed, and its reduction to and comparison with indentation mechanical model under a monotonic loading will be discussed. And then we will show theoretical analysis on a similarity of stress field between the rim of contact area between flat punch indenter and surface of a sample and crack tips. In section “[Indentation Fatigue Deformation](#),” extensive experimental results on polycrystalline copper and finite element analysis will be performed to demonstrate similarities between indentation fatigue depth propagation and fatigue crack growth, and an indentation fatigue depth propagation law will be developed. In section “[Indentation Fatigue Damage](#),” we will extend the indentation fatigue mechanics to indentation fatigue damage with a focus on indentation fatigue-induced failure of materials and predict the fatigue strength of materials from indentation fatigue. Concluding remarks will be given in section “[Concluding Remarks](#).”

Mechanics Theory of Indentation Fatigue

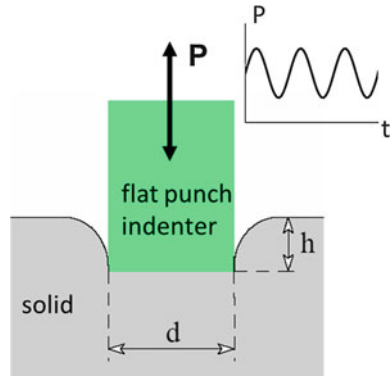
Indentation Load-Depth Curve

Consider a semi-infinite homogeneous elastic solid indented by a rigid frictionless flat punch, as schematized in Fig. 1, when the indentation load P is applied to the indenter, the penetration of the indenter (i.e., indentation depth) to the solid is h , and based on Hertz contact theory, it can be written as

$$h = \frac{(1 - \nu^2)}{Ed} P \quad (1)$$

where d is the diameter of the indenter, and E and ν are Young's modulus and Poisson's ratio of the solid, respectively. Equation 1 is also referred to as Sneddon's solution and serves the fundamental of determining Young's modulus of materials from a flat punch indentation (Sneddon 1965). Besides, because of the constant

Fig. 1 Schematic illustration of indentation fatigue on a semi-infinite solid via a flat punch indenter subjected to a fatigue loading P . d is the diameter of flat punch indenter, h is the indentation depth. The inset illustrates the cyclic fatigue loading



diameter of the punch indenter, Eq. 1 suggests a linear variation of indentation depth with indentation load at a small deformation, indicating one of the benefits of flat punch indenter in indentation technique in comparison with other indenter shapes such as sharp indenter, Vicker indenter, or spherical indenter that will result in increased contact area with the increase of the indentation depth. Obviously, when the indentation load P increases or decreases, Eq. 1 holds instantaneously for a semi-infinite homogeneous elastic solid, and thus it can be used for both loading and unloading conditions. Assume the indentation load P is applied with a sinusoidal loading spectrum, we will have

$$P(t) = P_m + \Delta P \sin(\omega t) \tag{2}$$

where ΔP , P_m , and ω are the indentation load range, mean, and frequency, respectively. With Eqs. 1 and 2, when a sinusoidal loading is applied to a semi-infinite homogeneous elastic solid via a flat punch indenter, the indentation load-depth relationship will be

$$h(t) = \frac{(1 - \nu^2)}{Ed} (P_m + \Delta P \sin(\omega t)) \tag{3}$$

That is

$$h(t) = \frac{(1 - \nu^2)}{Ed} P_m \left(1 + \frac{\Delta P}{P_m} \sin(\omega t) \right) \tag{4}$$

Apparently, the variation of indentation depth with experimental time is sinusoidal. Assume

$$h(t) = h_m + \Delta h \sin(\omega t) \tag{5}$$

where h_m and Δh are the mean and magnitude of the indentation depth. Thus, Eq. 4 can be rewritten as

$$h_m = \frac{(1 - \nu^2)}{Ed} P_m \quad (6)$$

$$\Delta h = \frac{(1 - \nu^2)}{Ed} \Delta P \quad (7)$$

Equation 6 is similar to Eq. 1 and indicates the indentation depth is induced by the mean load of the sinusoidal loading. Meanwhile, Eq. 7 indicates the dynamic response of indentation under the sinusoidal loading and can be written as

$$\frac{E}{1 - \nu^2} = \frac{1}{d} \frac{\Delta P}{\Delta h} \quad (8)$$

$E/(1 - \nu^2)$ is usually referred to as the stiffness of materials. Apparently one can easily probe the variation of stiffness of materials by applying a dynamic loading. Equation 8 has served the theoretical model in the measurement of dynamic contact stiffness of materials and has been employed in the continuous stiffness measurement (CSM) technique in the instrumented indentation technique (Li and Bhushan 2002; Asif et al. 1999; Loubet et al. 2000).

Indentation Stress Intensity Factor

Consider the solid in Fig. 2a, the boundary condition of stress field beneath the flat punch indenter is (Sneddon 1965; Johnson 1985)

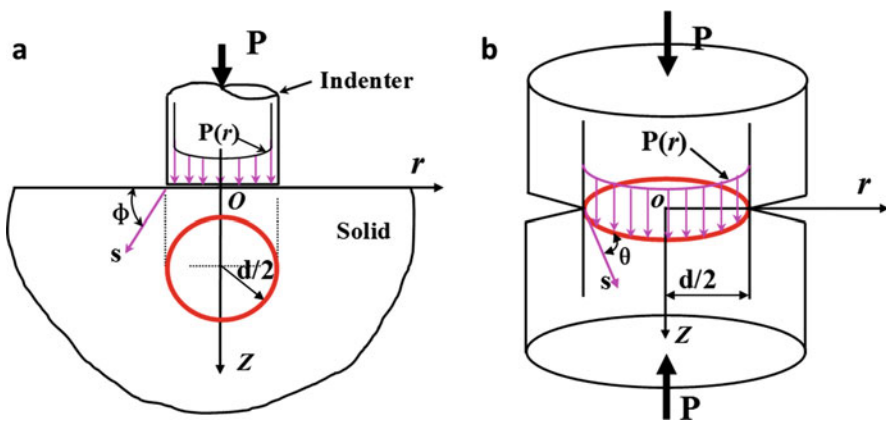


Fig. 2 (a) Schematic representation of contact between the flat punch indenter and a semi-infinite solid. (b) Circumferentially cracked notch specimen subjected to far-field compressive load Xu et al. 2009 (reprinted with the permission from Elsevier)

$$\sigma_{zz}(x, 0) = 0, \quad |x| > d/2 \quad (9a)$$

$$\tau_{xz}(x, 0) = 0, \quad 0 \leq |x| \leq d/2 \quad (9b)$$

In the cylindrical coordinate system r, θ, z , the stress field can be written as (Sneddon 1965; Johnson 1985)

$$\left\{ \begin{array}{l} \sigma_{zz} = \frac{Eh}{2(1-\nu^2)} \int_0^\infty (1+mz) e^{-mz} J_0(mr) \sin(md/2) dm \\ \sigma_{rr} = \frac{Eh}{2(1-\nu^2)} \left[\int_0^\infty (1-mz) e^{-mz} J_0(mr) \sin(md/2) dm \right. \\ \quad \left. - \frac{1}{r} \int_0^\infty \frac{1}{m} (1-2\nu-mz) e^{-mz} J_1(mr) \sin(md/2) dm \right] \\ \sigma_{\theta\theta} = \frac{Eh}{2(1-\nu^2)} \left[\int_0^\infty 2\nu m^2 e^{-mz} J_0(mr) \sin(md/2) dm \right. \\ \quad \left. + \frac{1}{r} \int_0^\infty \frac{1}{m} (1-2\nu-mz) e^{-mz} J_1(mr) \sin(md/2) dm \right] \\ \sigma_{zr} = \frac{Eh}{2(1-\nu^2)} \int_0^\infty mz e^{-mz} J_1(mr) \sin(mz) dm \end{array} \right. \quad (10)$$

where m is the variable function. J_α is the Bessel function of the first kind and is

$$J_\alpha(x) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m! \Gamma(m+\alpha+1)} \left(\frac{x}{2}\right)^{2m+\alpha}$$

Within the contact range at $z = 0$, one can have

$$\sigma_{zz}|_{z=0} = \frac{Et_0}{\pi(1-\nu^2)} \int_0^\infty J_0(mr) \sin(md/2) dm \quad (11)$$

Because $\int_0^\infty J_0(mr) \sin(md/2) dm = \frac{1}{\sqrt{(d/2)^2 - r^2}}$, $r \leq d/2$, one will have

$$\begin{aligned} \sigma_{zz}|_{z=0} &= \frac{Et_0}{\pi(1-\nu^2)} \int_0^\infty J_0(mr) \sin(md/2) dm \\ &= \frac{Et_0}{\pi(1-\nu^2)} \frac{1}{\sqrt{(d/2)^2 - r^2}} \quad r \leq d/2 \end{aligned} \quad (12)$$

That is

$$\sigma_{zz}(r, 0) = \frac{Eh}{\pi(1-\nu^2)} \frac{1}{\sqrt{(d/2)^2 - r^2}}, \quad r \leq d/2 \quad (13)$$

At the rim of indentation, $|x|$ is approaching to $d/2$, one will have

$$\sigma_{zz}(r, 0) = \frac{Eh}{\pi(1-\nu^2)} \left(\frac{1}{\sqrt{dr}} + \dots \right), \quad r \rightarrow d/2 \tag{14}$$

Equation 14 indicates there is a stress singularity at the edge of indentation with respect to \sqrt{r} .

Similarly, for a semi-infinite model-I crack subjected to a far-field compressive load P , as schematized in Fig. 2b, given the distanced $d/2$ from the origin of the coordinate system to the crack tip, the stress field at the crack tip is (Hertzberg 1995)

$$\begin{pmatrix} \sigma_{zz} \\ \sigma_{rz} \\ \sigma_{rr} \\ \sigma_{\varphi\varphi} \end{pmatrix} \rightarrow -\frac{2P}{\pi d^2 \sqrt{s}} \cos \frac{\varphi}{2} \begin{pmatrix} 1 + \sin \frac{\varphi}{2} \sin \frac{3\varphi}{2} \\ \sin \frac{\varphi}{2} \cos \frac{3\varphi}{2} \\ 1 - \sin \frac{\varphi}{2} \sin \frac{3\varphi}{2} \\ 2\nu \end{pmatrix} \tag{15}$$

Comparison between Eqs. 14 and 15 implies that σ_{zz} in Eq. 15 is an asymptote of Eq. 13 at $r \rightarrow d/2$, leading to an equivalence between Eqs. 14 and 15, and thus suggesting the same stress singularity with respect to $\sqrt{}$ at the rim of flat punch indentation and crack tip.

Given Eq. 1, Eq. 13 can be rewritten as

$$\sigma_{zz}(r, 0) = -\frac{4P}{\pi d^2}, \quad r \leq d/2 \tag{16}$$

According to the definition of stress intensity factor, we will have

$$\sigma_{zz} = -\frac{K_I}{\sqrt{2\pi(d/2-r)}} \tag{17}$$

And thus the stress intensity factor near the contact rim of flat punch and solid can be expressed as

$$K_I = \frac{F}{d\sqrt{\pi d/2}} \tag{18}$$

And Eq. 18 indicates the same stress intensity factor near the contact rim of flat punch and solid with that at the crack tip.

Under a fatigue loading condition, the stress intensity factor can be obtained similar to that of Eq. 2, and is

$$K_{\max} = \frac{P_{\max}}{d \sqrt{\pi d/2}}, \quad \Delta K = \frac{\Delta P}{d \sqrt{\pi d/2}} \quad (19)$$

where $P_{\max} = \Delta P/2 + P_m$ is the maximum indentation load.

Indentation Fatigue Deformation

Indentation Fatigue Depth Propagation Law

When a solid (engineering metals otherwise specified) is subjected to a fatigue loading, crack will initiate and propagate with the increase of cyclic number. The steady state (i.e., the 2nd stage) of the fatigue crack growth $(da/dN)_s$ can be described by the well-known Paris equation (Ritchie 1977)

$$(da/dN)_s = C (\Delta K)^q \quad (20)$$

where $\Delta K = K_{\max} - K_{\min}$ is the nominal stress intensity factor range, and K_{\max} and K_{\min} are the maximum and minimum stress intensity factors, respectively. To highlight the competition between the intrinsic mechanism of crack tip growth and extrinsic mechanism of crack-tip shielding behind the tip, Eq. 20 can be modified to (Liu and Chen 1991; Dauskardt et al. 1992)

$$(da/dN)_s = C (K_{\max})^n (\Delta K)^m \quad (21)$$

where ΔK and K_{\max} describe the intrinsic and extrinsic mechanisms of fatigue crack growth, respectively, and their dominance is reflected by the power indices m and n . m , n , and C are empirical constants and depend on material and microstructure, fatigue frequency, loading mode and environment, etc.

When an elastoplastic solid is subjected to a fatigue loading via a flat punch indenter, given the similarity of stress singularity in theory in section “[Indentation Stress Intensity Factor](#)” between the crack tip and the rim of indentation, in particular, the same stress intensity factor near the rim of indentation and crack tip, and inspired by Eq. 21, the indentation fatigue depth propagation at the steady state can be described quantitatively by using a power law

$$(dh/dN)_s = C_i K_{\max}^{n_i} \Delta K^{m_i} \quad (22)$$

where ΔK and K_{\max} are obtained from Eq. 19. C_i , n_i , and m_i are constants of indentation fatigue and depend on material/microstructure and testing environments. Similar to the fatigue crack growth, the continuous sinking of the indenter into the solid suggests the accumulation of plastic deformation beneath the indenter and is driven by the stress concentration near the rim of contact between the indenter and solid surface. It reflects the intrinsic mechanism of indentation depth propagation

with the cyclic number and is determined by ΔK . In contrast, K_{\max} indicates the elastic recovery during the indentation depth propagation and represents the extrinsic mechanism. We will validate Eq. 22 through both computations and experiments in the following.

Computational Validation

Computational Method and Modeling

Modeling plastic behavior of materials under a fatigue loading has received considerable attention, and it requires the including of the Bauschinger effect, elastic shakedown, cyclic hardening or softening, ratcheting, and mean stress relaxation. A significant number of models have been developed over the last decades such as developed by Chaboche and Nouailhas (1989a, b), Ohno and Wang (1995), and Delobelle et al. (1995), and most of these models are based on the kinematic hardening model of Armstrong and Frederick (A-F model) (Jiang and Kurath 1996). A-F model has also been employed in the study of cyclic spherical indentation behavior by Huber and Tsakmakis (Moosbrugger and Morrison 1997; Abdel-Karim and Ohno 2000), but limited to a few number of cycles. In our simulations, two different polycrystalline copper alloys whose elastoplastic property proves to obey the classical kinematic hardening rule (namely the A-F model) will also be employed. The mechanical properties for a brittle copper alloy are 122.5 GPa for Young's modulus, 0.35 for Poisson's ratio, 33.32 MPa for initial yield stress, and 1.607 GPa for a linear hardening rate; the mechanical properties for a ductile copper alloy are 119.9 GPa for Young's modulus, 0.35 for Poisson's ratio, 73.50 MPa for initial yield stress, and 0.369 GPa for a linear hardening rate (Mclean 1965). We note that these two copper alloys can be distinguished based on measurements from the conventional uniaxial fatigue cracking test. For example, for fatigue cracking behavior of the brittle copper alloy, the index n is higher than m in Eq. 21, and the index n is less than m in Eq. 21 for fatigue cracking behavior of the ductile copper alloy. The plat punch indenter with radius $d/2 = 1.0$ mm is assumed to be rigid and frictionless. Figure 3a illustrates the history of fatigue loading with a sinusoidal manner. The application of a preloading to the mean fatigue load will help to minimize the effect of initial loadings. The simulations were conducted by the finite element software ABAQUS.

Results and Discussion

For the brittle copper as a representative, under $\Delta P = 100$ MPa and $P_m = 100$ MPa, the indentation load (P)-depth (h) curve is shown in Fig. 3b, and three stages are observed in response to the loading history in Fig. 3a. Stage II corresponds to the cyclic response of indentation to the fatigue loading, and is our focus. At the beginning of this stage, the indentation depth propagation rate (per cycle) is quite high, and then it decays with the increase of number of cyclic loading till to approximately reaching a constant. Overall, the indentation depth keeps increasing upon cyclic loading and shows a clear difference from the static elastoplastic

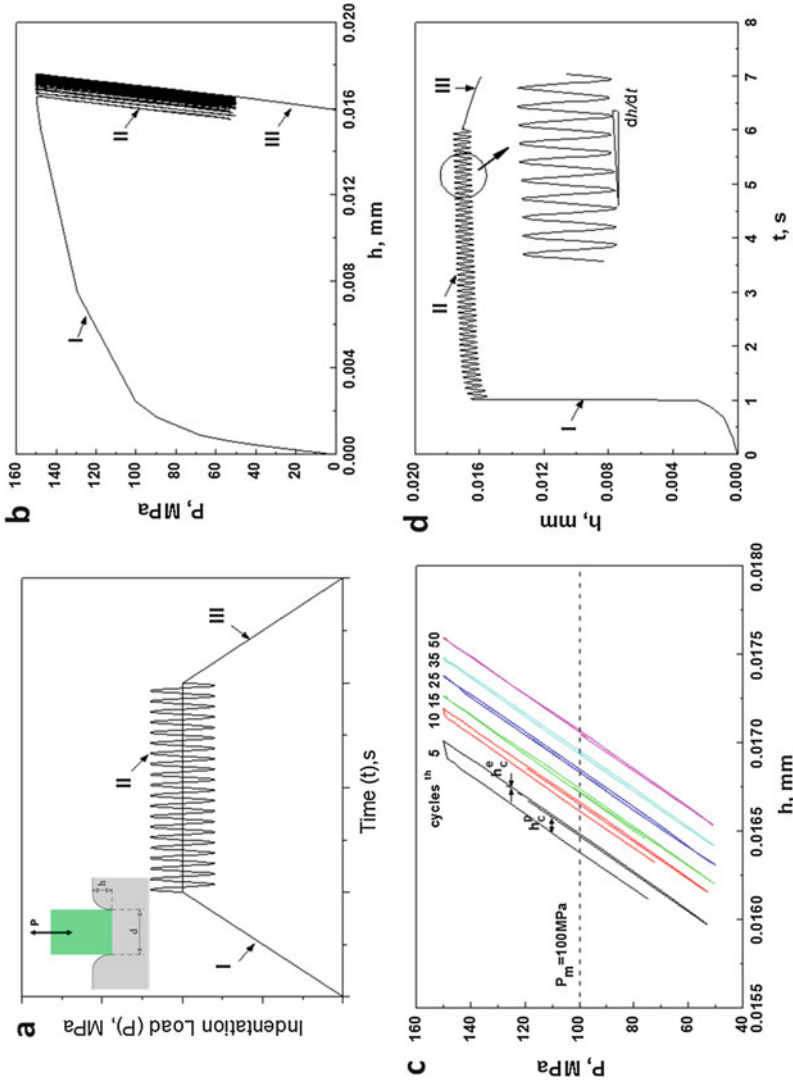


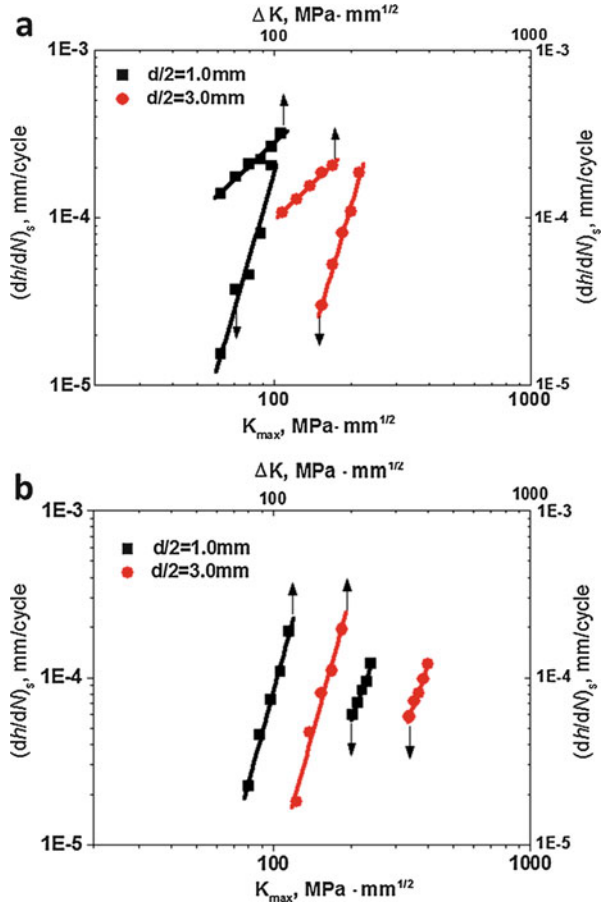
Fig. 3 (a) Schematic of the load history used in the numerical simulation of indentation fatigue. (b) The typical indentation load (P)-depth (h) curve. (c) The hysteresis loop of indentation load (P)-depth (h) curves under fatigue loading cycles 5th, 10th, 15th, 25th, 35th, and 50th. (d) Indentation depth (h)-time (t) curves. $\Delta P = 100$ MPa and $P_m = 100$ MPa Xu et al. 2009 (reprinted with the permission from Cambridge University Press)

indentation response. Besides, the propagation of indentation depth shows the same sinusoidal frequency as that of the applied indentation load. More importantly, there is a hysteresis loop in the indentation fatigue P - h curve at each cyclic loading, and several of them are magnified in Fig. 3c. As the number of cycles increases, the hysteresis loop tends to be more “closed” in comparison with the early ones, leading to a smaller enclosed area. Further, each hysteresis loop can be decomposed into two parts: the cyclic elastic indentation depth range h_c^e and the cyclic plastic indentation depth range h_c^p . The former suggests the recovery of indentation depth during indentation fatigue, and the latter is associated with the continuous increase of indentation depth due to the propagation of plastic zone beneath the indenter. Both h_c^e and h_c^p decrease with the increase of indentation depth toward the overall closed hysteresis loop, and the decreasing rate for h_c^p seems to be larger. Essentially, the area of hysteresis loop reflects the dissipated energy that is required to advance the indentation depth, and the dissipation rate will slow down as the indentation depth propagation increases (Xu et al. 2007b).

Figure 3d presents the variation of indentation depth with the history of cyclic loadings. Similar to the indentation fatigue P - h curve in Fig. 3b, three stages that correspond to the loading history are also obtained, and the second stage under a cyclic loading will be investigated. It further confirms that an indentation depth rate (dh/dN) arrives in a stable state with the increase of cyclic numbers at the second loading stage, and this stable indentation depth propagation ratio $(dh/dN)_s$ depends on the indentation load range and the maximum load, and radius of indenter, i.e., ΔK and K_{\max} and $d/2$. With $d/2 = 1.0$ m, Fig. 4a gives the variation of the steady-state depth rate $(dh/dN)_s$ with ΔK and K_{\max} . A power-law function (i.e., Eq. 22) of $(dh/dN)_s$ as both the stress intensity range ΔK and the maximum stress intensity K_{\max} is obtained. When the indenter radius $d/2 = 3.0$ mm, the functional form of Eq. 22 holds. Further, the exponents n_i , m_i can be obtained, and $n_i = 5.2 > m_i = 1.5$, suggesting that $(dh/dN)_s$ is dominated by K_{\max} in comparison with ΔK . Similar to the fatigue crack propagation, ΔK can be considered the driving force for propagation of plastic zone due to the stress concentration at the rim of contact (associated with the cyclic plasticity), which promotes indentation depth propagation and represents the intrinsic mechanism. Meanwhile, K_{\max} is associated with the elastic recovery of indentation depth, which recovers elastic deformation during indentation fatigue and represents the extrinsic mechanism. The dominance of K_{\max} in the indentation fatigue depth propagation for the brittle copper alloy is qualitatively similar to the conventional fatigue cracking behavior of brittle materials (Liu and Chen 1991; Dauskardt et al. 1992; Ritchie 1999), where their crack growth rate is dominated by K_{\max} in Eq. 21.

Following the similar computational procedures, we performed the finite element analysis of indentation fatigue on the ductile copper alloys. As ΔK and K_{\max} change with different indenter radii, Fig. 4b shows the corresponding obtained $(dh/dN)_s$, which confirms that the indentation fatigue depth propagation law (Eq. 22) holds. The exponents are $n_i = 4.2 < m_i = 5.6$, which indicates that the dominance of ΔK over that of K_{\max} for ductile copper alloys, also consistent well with the fatigue cracking behaviors of ductile materials (Liu and Chen 1991; Dauskardt et al. 1992).

Fig. 4 Variation of the steady-state rate of indentation fatigue depth $(dh/dN)_s$ with the stress intensity range ΔK and the maximum stress intensity K_{max} from the numerical simulation on (a) a brittle copper alloy and (b) a ductile copper alloy Xu et al. 2009 (reprinted with the permission from Elsevier)



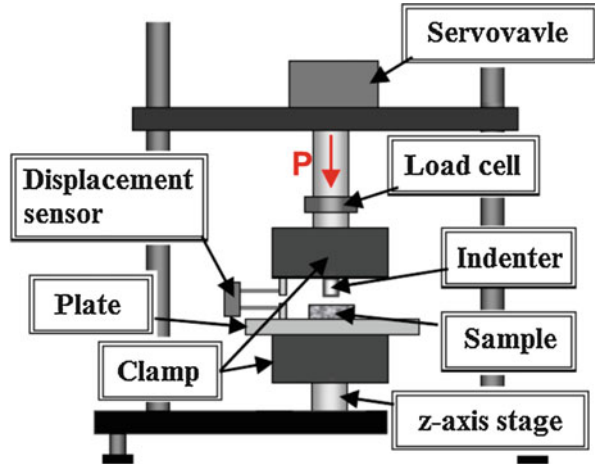
In summary, these computational results validate the proposed power-law function, i.e., Eq. 22, and the steady-state indentation fatigue depth propagation rate is a power-law function of K_{max} and ΔK for both brittle and ductile metals, similar to that fatigue crack growth. Besides, the dominance of K_{max} and ΔK can be determined by comparing their exponents. In addition, note that the coefficients in Eq. 22 C_i , n_i , and m_i depend on materials and testing conditions but are insensitive to indenter radii.

Experimental Validation

Material Choice and Testing Platform

99.9% polycrystalline copper (with 0.0262% Zn, 0.0145% P, 0.003% Pb, and 0.1266% Fe) with an average grain size $32\mu\text{m}$ was chosen, and the surface of each specimen (15 mm in diameter, by 15 mm in length) was polished to minimize the

Fig. 5 Schematic of indentation fatigue testing system



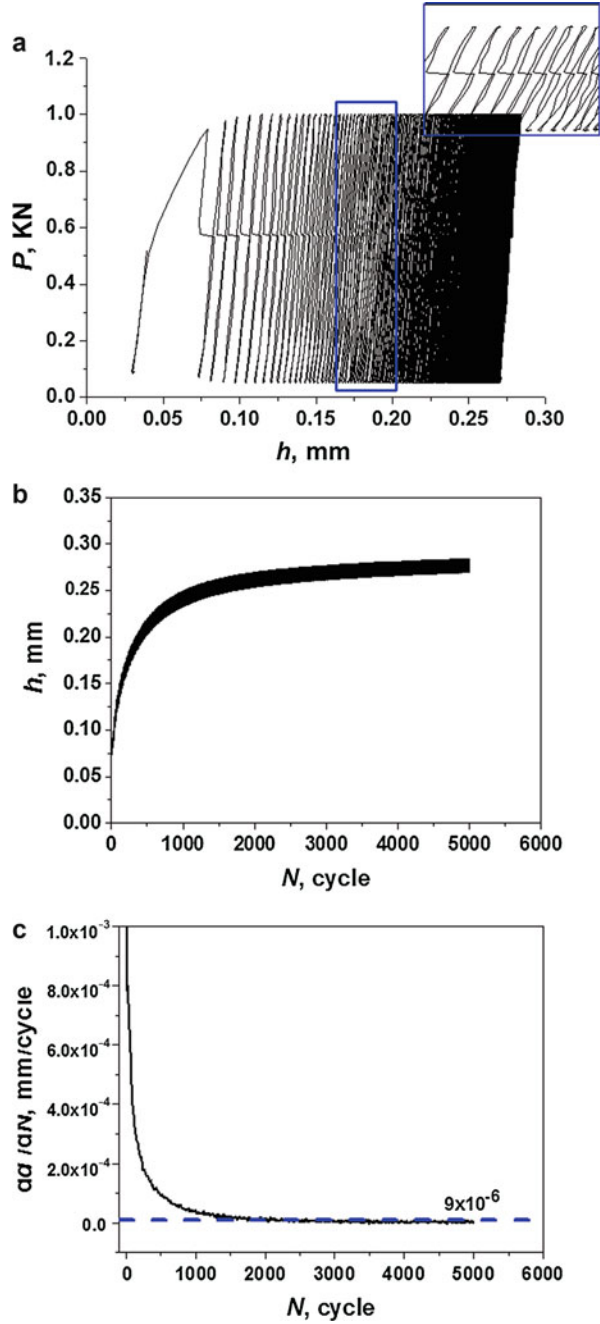
effect of surface roughness before each experiment. Besides, annealing procedures for specimens were made at temperatures $150\text{ }^{\circ}\text{C}$ for 35 minutes to relieve the residual stress. The material of the flat punch indenter was made by high temperature alloys and its deformation was neglected in comparison with the indentation depth. The diameter of the indenter was 0.993 mm .

A sinusoidal wave was employed to apply the indenter and was used to mimic a fatigue loading condition and the loading frequency ($f=\omega/2\pi$, Eq. 2) was chosen to be 1 Hz . To keep contact between the indenter and surface of the specimens during the experiment, the minimum indentation load (P_{min}) was set to 20 N . The conventional tensile machine INSTRON8871 that can easily realize the fatigue loading has been designed to conduct the indentation fatigue experiment with a customized load cell, as schematized in Fig. 5. All experiments were performed at room temperature, and the indentation load-depth data were recorded automatically for each 10 loading cycles during an indentation fatigue testing.

Experimental Results

For representative fatigue loadings, $P_{max} = 200\text{ N}$ and $P_{min} = 20\text{ N}$ were employed and the corresponding $K_{max} = 485.26\text{MPa} \cdot \text{mm}^{1/2}$ and $\Delta K = 469.12\text{MPa} \cdot \text{mm}^{1/2}$ can be calculated by utilizing Eq. 22. Figure 6a shows the recorded indentation load-depth curve. The indentation depth increases with the cyclic number. Besides, there is a quick penetration of the indenter at the initial stage, and then a relative steady state of indentation penetration rate arrives. More importantly, a hysteresis loop is observed in each loading-unloading cycle and seems to remain unclosed even after 3000 cycles. Figure 6b presents the curve of the indentation depth-loading cycle, where the width of the bands is expected to associate with the elastic recovery in each loading cycle. Similar to that observation in simulations in Fig. 3d, the indentation depth propagation shows a quick increase at the beginning and then followed by a steady-state stage. Essentially, these features can be understood from

Fig. 6 Indentation fatigue experiment on polycrystalline copper. (a) Indentation load P - indentation depth h . (b) Indentation depth h and (c) indentation depth per cycle dh/dN with the applied fatigue cycle number. $K_{max} = 485.26\text{MPa} \cdot \text{mm}^{1/2}$ and $\Delta K = 469.12\text{MPa} \cdot \text{mm}^{1/2}$ Xu and Yue 2006 (reprinted with the permission from Cambridge University Press)



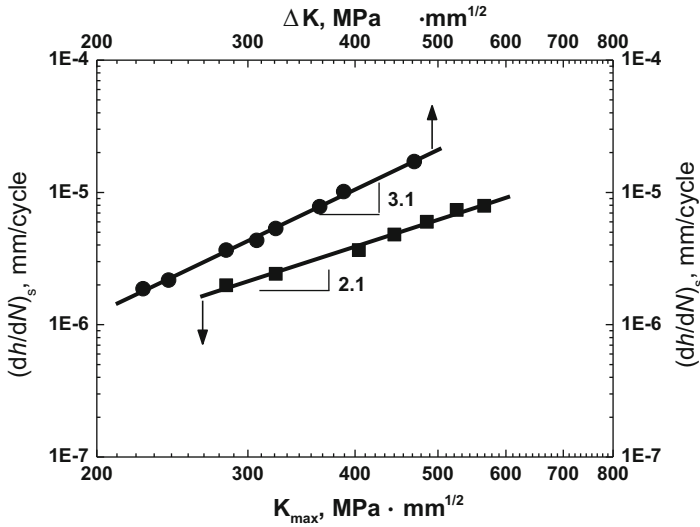


Fig. 7 Variation of the steady-state rate of indentation fatigue depth $(dh/dN)_s$ with the stress intensity range ΔK and maximum stress intensity K_{\max} , measured from experiment on a polycrystalline copper Xu et al. 2009 (reprinted with the permission from Elsevier)

the dislocation activation and retraction, and more details see section “[Discussion on Deformation Mechanism](#).” In fact, these features are quite analogous to the first two stages of uniaxial tensile fatigue crack growth. The approximate steady-state propagation of indentation fatigue is further confirmed when the indentation fatigue depth propagation rate dh/dN is plotted as a function of number of cycles in Fig. 6c, and the steady-state propagation rate is $(dh/dN)_s$ as denoted in Eq. 22.

Given $(dh/dN)_s$, we can examine its variation with ΔK and K_{\max} and validate the indentation fatigue depth propagation law (i.e., Eq. 22). Figure 7 shows there is a linear variation of $(dh/dN)_s$ with both ΔK and K_{\max} on log-scale plots, which agrees with Eq. 22. Besides, the power exponents are $m_i = 3.1$ and $n_i = 2.1$, respectively. The higher m_i implies that the steady-state propagation of indentation fatigue is dominated by ΔK , which is also consistent with fatigue crack growth in ductile metals (Ritchie 1999). We should note the power exponents n_i and m_i may not exactly equal to the n and m obtained from uniaxial tensile fatigue crack propagation because of influences of local microstructures in indentation experiments, and this difference is expected to increase with the decrease of indentation scales.

Extend to Overloading and Underloading

In previous sections, we have revealed the similarity between indentation fatigue depth propagation and fatigue crack growth under a normal fatigue loading spectrum. It is known that the fatigue crack growth is greatly affected by the applied

fatigue loading interactions such as sudden increase of peak load followed by a normal fatigue loading conditions (referred to as overloading here) and sudden decrease of peak load followed by a normal fatigue loading conditions (referred to as underloading here). For example, the normal fatigue crack growth can be delayed and boosted by the overloading and underloading, respectively. In this section, we will continue to employ polycrystalline copper to investigate whether the overloading/underloading will lead to delay/acceleration of indentation fatigue depth propagation so as to further study its similarity to fatigue crack growth.

Figure 8a shows the indentation fatigue P - h curve under an overloading condition, and the inset illustrates the overloading spectrum, where two loading blocks with a low-high loading sequence are considered. P_{max0} and P_{min} are the maximum load and the minimum load of the first load block, respectively. P_{max1} is the maximum load of the second load block. N_0 and N_1 are the number of loading cycles of the two load blocks, respectively. The indentation depth increases with the number of cycles, and its propagation arrives at a relatively steady stage with a constant rate after an initial sharp increase in each loading block. Besides, hysteresis loops are also found in each block and seem not be closed, even after 7000 cycles in the second loading block.

The indentation depth-loading cycle is plotted in Fig. 8b, and it further indicates the approximately constant evolution of indentation depth propagation after initial rapid adjustment in each loading block, similar to that under a single normal fatigue loading block. Figure 8b further shows the rate of indentation depth propagation dh/dN with the number of fatigue loading. In each block, a clear steady stage with a constant dh/dN is observed after a quick decrease. More importantly, when the maximum load increases from 500 N in the first loading block to 700 N in the second loading block, a dh/dN shows an obvious increase first and then gradually decreases until to a new higher steady state. Experiments further indicate that the new steady state is nearly independent of the maximum loads in the second load block, except that it needs more cycles to reach the new steady state for the larger maximum load, as shown in Fig. 8b, where the maximum load changes from 700 N to 800 N in the second load block. Generally, a larger increase of the maximum load will require more number of cycles to eliminate the effect of load interaction. The enhancement of the indentation depth propagation by the overloading further shows similarity with that in conventional fatigue crack propagation (Tvergaard 2005; Sadananda and Vasudevan 2003; Huang and Ho 2000, 2003; Borrego et al. 2003; Kumar et al. 1996).

Similar to overloading conditions, the underloading condition also consists of two fatigue loading blocks with a decrease of the maximum load, i.e., a high-low loading sequence in the two load blocks, as illustrated in the inset in Fig. 8c. Consider a decrease of the maximum load from $P_{max0} = 600$ N to $P_{max1} = 550$ N with $N_0 = 2000$ cycles, and $N_1 = 7000$ cycles, Fig. 8c shows the indentation depth - number of cycles. The indentation depth keeps increasing instead of ceasing when the maximum load decreases. Besides, similar to that under overloading conditions, dh/dN arrive at a new steady state after lowering the maximum load, yet within a very short cyclic numbers, as shown in Fig. 8d. The steady state dh/dN for the new

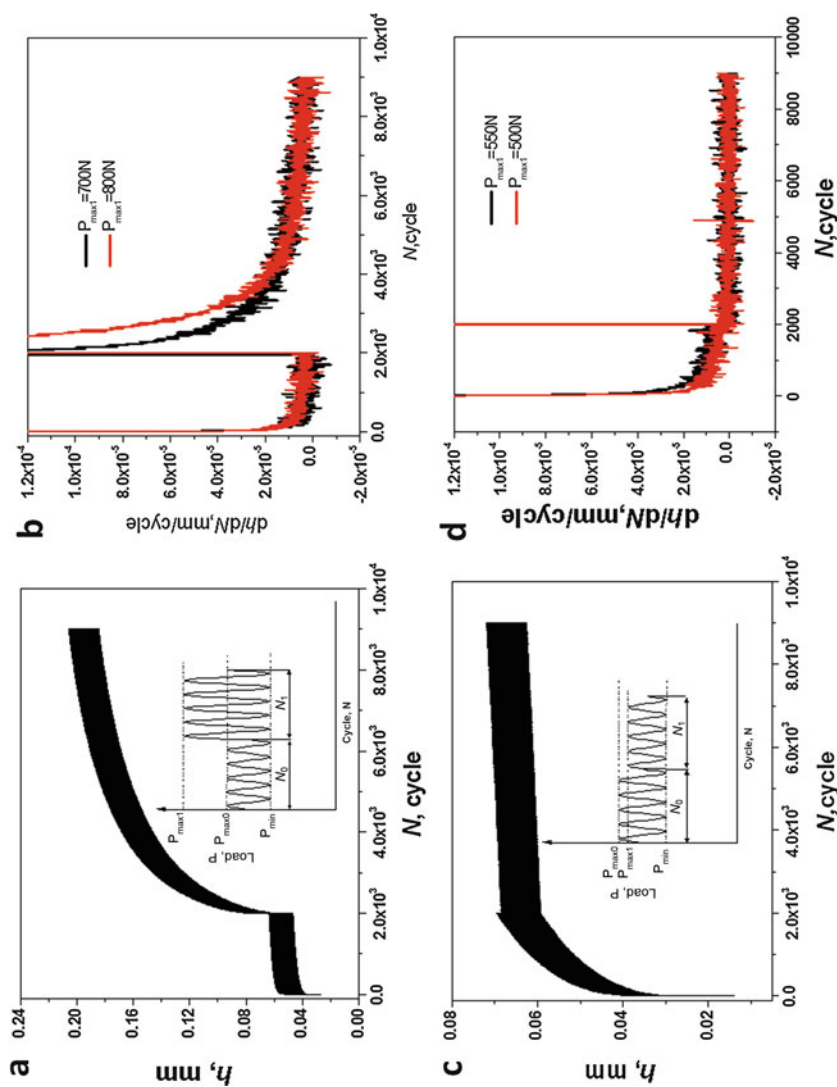


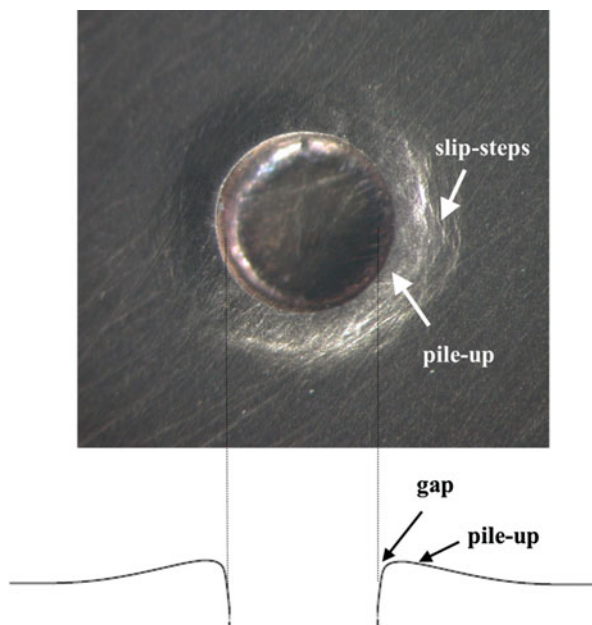
Fig. 8 (a) Indentation depth and (b) indentation depth per cycle - number of cycles curves under a low-high overloading (inset in a). $P_{max0} = 500 N$, $P_{maxI} = 700 N$, $N_0 = 2000$ cycles, $N_I = 7000$ cycles. (c) Indentation depth and (d) indentation depth per cycle - number of cycles curves under a high-low underloading (inset in c). $P_{max0} = 600 N$, $P_{maxI} = 550 N$, $N_0 = 2000$ cycles, $N_I = 7000$ cycles Xu et al. 2007a (reprinted with the permission from Elsevier)

maximum load is obtained in the second cyclic loading block and smaller than that for the first cyclic loading block, indicating a delayed propagation of indentation depth. When the decrease of the maximum load is much larger (from $P_{max0} = 600$ N to $P_{max1} = 500$ N), Fig. 8d shows that the indentation depth propagation becomes slower. The delay of indentation fatigue depth propagation due to underloading is also similar to those findings in conventional fatigue crack propagation.

Microstructural Observation

Figure 9 presents the optical image of indentation on polycrystalline copper after cyclic loadings (Fig. 8c for indentation fatigue load-depth curve), where $P_{max0} = 600$ N, $P_{max1} = 550$ N, $N_0 = 2000$ cycles, $N_1 = 7000$ cycles. Pile-up, gap, and wrinkles (i.e., overlapping layers) are observed, similar to those under monotonous loading conditions. Generally, the pile-up or sink-in is caused by material hardening and when the hardening exponent of materials, n , is less than 0.33 (Storakers and Larsson 1994; Taljat and Pharr 2004), the pile-up will appear near the flat punch indentation (Storakers and Larsson 1994). For the polycrystalline copper used in our experiment, $n \approx 0.3 < 0.33$ was measured from our uniaxial tensile testing (Xu et al. 2006), and its cyclic hardening exponent $n \approx 0.1$ is also less than 0.33, which agrees well with the appearance of pile-up. Given the constant contact area between the flat punch indenter and surface of materials, deformation of materials will deviate from the indenter in accommodation with the pile-up,

Fig. 9 Optical microscope (OM) indentation image of polycrystalline copper (top view) and its sketch map (slide view) Xu et al. 2007a (reprinted with the permission from Elsevier)



leaving a gap between pile-up and indentation. This gap will increase with the continuous sinking of indenter into the materials, which also agrees with the finite element analysis (Xu and Yue 2007). Wrinkles reflect the slip-steps of shear bands surrounding the indentation, and the formation and propagation of these shear bands can be deemed the major plastic deformation mechanism in the present indentation fatigue on polycrystalline copper. The closer to the indentation, the more wrinkles due to the more severe plastic deformation. From crystal plasticity theory point of view, the wrinkles result from activation of multiple slip systems, and these wrinkles are also referred to multislip-steps in quasi-static indentation (Lloyd et al. 2005; Zaafarani and Raabe 2008; Nibur and Bahr 2003).

Figure 10 shows the SEM image of cross section of indentation. Pile-up and gap are observed more clearly. Besides, some light and dark cyclic lines with radii that decrease with increasing distance from the indenter are also observed. Those lines are expected to be dependent of the slip lines beneath the flat punch indenter, which is considered major deformation mechanism (Hill 1998). Figure 10 (i–iii) shows the local deformation near the indentation, highlighted in the boxes. Figure 10i shows the SEM images of pile-up around the samples. The pile-up is flat at first and then

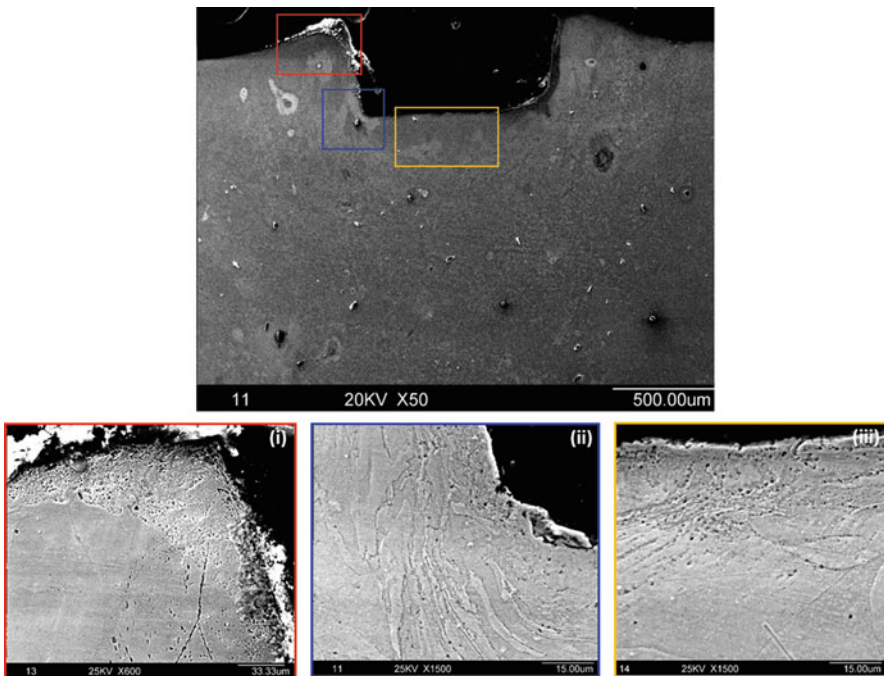


Fig. 10 SEM image of the cross section of indented polycrystalline copper sample under a high-low underloading (inset in Fig. 8c) and higher magnification (i) around the pile-up on the surface, (ii) near the corner of indentation, (iii) below the indentation; $P_{max0} = 600\text{ N}$, $P_{max1} = 550\text{ N}$, $N_0 = 2000\text{ cycles}$, $N_1 = 7000\text{ cycles}$ Xu et al. 2007a (reprinted with the permission from Elsevier)

vanishes with the distance far from the indentation. The flat transition is expected to be related with the amount of plastic deformation of materials around the indenter, and a cavity layer or cavity strip is similar to the form of a pile-up in geometry. From Fig. 10i, many cavities appear along the cyclic lines, and even some cracks in higher density cyclic lines near the rim of indentation. These cavities are expected to be nucleated at an early stage of indentation fatigue and accumulate to cracks, and then further spread with the increase of the number of loading cycles. Once the cracks interact, peel-offs of materials will happen, especially near the rim of indentation with higher cyclic lines associated with a high stress concentration. Figure 10iii gives the SEM images in the regions below the indentation. More cavities and cracks are found at a distance of about $\sim 15\mu\text{m}$ far from the free end of indentation. The orientation of these cracks is in line with the dark and light cyclic lines, which are approximately level near the end of the indentation and steeper further from the end of indentation.

Discussion on Deformation Mechanism

In the viewpoint of indentation load-indentation depth curves, the increase of indentation depth results from deformation ratcheting behavior of materials during cyclic indentation due to a nonzero mean fatigue stress. The ratcheting behavior reflects the accumulation of plastic deformation preceding in one direction. Besides, the nucleation and development of the cracks further facilitate the development of plastic zone, leading to a continuous increase of indentation depth.

When a new loading block is applied following a normal fatigue loading history, the retardation or acceleration of indentation fatigue depth propagation is caused by interactions of loading blocks and can be understood by considering activation of dislocations. For example, the effective applied stress σ_{eff} acting on mobile dislocation is

$$\sigma_{eff} = \sigma_{app} - \sigma_r \quad (23)$$

Where σ_{app} is the maximum fatigue stress and σ_r is the residual stress. For example, when the maximum load increases (i.e., overloading), the effective applied stress σ_{eff} is very large, meanwhile the residual stress σ_r is small, leading to a high indentation depth per cycle. At the same time, this difference will require numbers of loading cycles to increase the residual stress σ_r until reaching a new balance with a new steady state of indentation depth propagation. Hence, the greater the overloading amplitude, the more number of cycles are needed to remove the effect of overloading, consistent with experimental results in Fig. 8b. Besides, with the increase of the effective applied stress σ_{eff} , the dislocation density becomes intensive, the initial indentation depth per cycle will be very large and then gradually reduces and approaches to a new steady-state value with the arrival of new balance. The variation of the residual stress σ_r will lead to rearrangement of the dislocations beneath the indenter. The larger increase of the maximum load, the more cyclic

numbers of loading to rearrange dislocation, and the larger steady indentation depth propagation ratio. These interpretations further show the similarity with the generation and annihilation of dislocations around the crack tip under the fatigue loadings.

Similarly, the decrease in the maximum cyclic load after reaching steady indentation depth ratio will reduce the residual stress σ_r . Consequently, unloading with a smaller maximum load produces a smaller effective stress σ_{eff} for plastic zone propagation, leading to a decreased indentation depth propagation after initial adjustments, as shown in Fig. 8d, which is also similar to that of fatigue crack growth under fatigue loadings.

Indentation Fatigue Damage

Indentation Fatigue Strength Law

With the continuous increase in indentation load cycles, damage of indented materials will nucleate and propagate, and eventually leads to failure of materials (Lawn and Wilshaw 1975; Xu et al. 2007a; Guiberteau et al. 1993; Bhowmick et al. 2007). Consider the number of indentation cycles to failure as N_f and the amplitude of the cyclic indentation load as F_a , in this section, we will build the relationship between N_f and F_a .

For a uniaxial fatigue testing, the relationship between the amplitude of fatigue load, σ_a , and the number of cycles to failure of materials, N_f , is the well-known S-N curve and can be expressed as a power-law relationship (Suresh 1998; Basquin 1910)

$$\sigma_a = \sigma_f (N_f)^n \quad (24)$$

where σ_f and n are the fatigue strength coefficient and exponent, respectively, and are related to the material/microstructure and testing environment/condition. In general, σ_f equals the uniaxial fracture strength of the material under quasi-static loading.

Inspired by this equation (Eq. 24), we propose that a similar power-law relationship exists for indentation fatigue damage

$$F_a = F_f (N_f)^m \quad (25)$$

where F_f and m are the indentation fatigue strength coefficient and exponent, respectively, and both depend on the material/microstructure and testing environment/condition. In the following, we will validate this power law equation in experiments.

Experimental Validation

Material Choice and Testing Platform

Different from the employment of copper alloy, to highlight the damage, relatively brittle materials were used in this section, and they were PVC bulk material, and TiN, NiP thin films on steel (SUS304) substrate. Besides, in order to avoid the early damage by stress concentration at contact, a spherical indenter was employed (instead of a flat punch indenter or a pyramidal/conical indenter). The cyclic indentation load applied on the indenter varies in a sinusoidal manner; the load ratio, i.e., the ratio between the minimum and maximum indentation loads, was fixed at 0.1 for the thin-film specimens and fixed at 0.05 for the PVC bulk specimen. Similar to the study of indentation fatigue depth propagation on polycrystalline copper, the minimum load was above zero to keep the “compressive” contact all the time. The frequency of load is 20 Hz for thin films and 5 Hz for the bulk material. The maximum indentation force is constrained so as not to induce cracking under static/monotonic loading conditions.

The failure was monitored in situ by acoustic emission (AE) during the indentation fatigue, as schematized in Fig. 11. This technique was used to diagnose the nucleation and propagation of damage and cracks in the static and cyclic indentation tests and showed good agreement with the observations of microstructures (Yonezu et al. 2009, 2010). Since AE may also be sensitive to contact friction/noise due to cyclic indentation, the initiation of major crack, or material failure, is quantitatively defined if more than 20 new AE counts per second are detected. This ensured the detection of the initiation of major cracks (macrocracks), which is regarded as failure in this study. From the variation of the AE count, the critical number of fatigue cycles (upon failure) can be readily determined, which is regarded as the fatigue strength of the specimen under a particular testing condition.

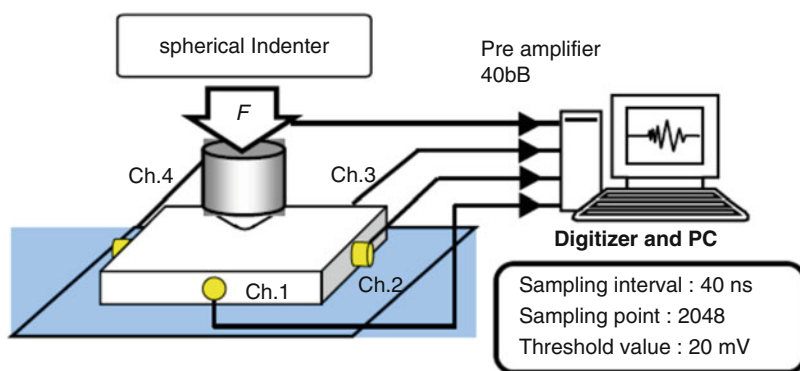


Fig. 11 Experimental setup of spherical indentation with AE and corrosion potential fluctuation monitoring system

Results and Discussion

Indentation fatigue tests on the PVC bulk material were first performed. Figure 12a shows a typical curve of indentation fatigue depth versus the number of cycles with a maximum indentation force of 880 N. Both the maximum and minimum depth histories are shown and the difference between them is due to the elastic recovery. The indentation depth increases with the number of cycles, and after an initial stage, a steady-state indentation depth propagation rate is achieved. These features are similar to those of polycrystalline copper in section “[Experimental Validation.](#)” The accompanied AE count shows that at $N = 109,245$ cycles there is a significant increase of new AE counts. This implies that a macrocrack nucleates around the indentation, which is identified as the failure in the PVC, as shown in the inset in Fig. 12a. For the same specimen, more indentation fatigue tests with different load amplitudes were carried out, and the resulting number of cycles to failure obtained as a function of the fatigue load amplitude. This is shown in Fig. 12b. A linear

Fig. 12 (a) Experimental evolution of indentation fatigue depth and AE count with number of cycles on the PVC bulk material. The inset shows the surface topography near the indentation (after failure). (b) The variation of the indentation fatigue load amplitude with number of cycles to failure for the PVC bulk material. The maximum indentation fatigue load is 880 N and the load ratio is 0.05. The frequency is 5 Hz Xu et al. 2010 (reprinted with the permission from Taylor & Francis)

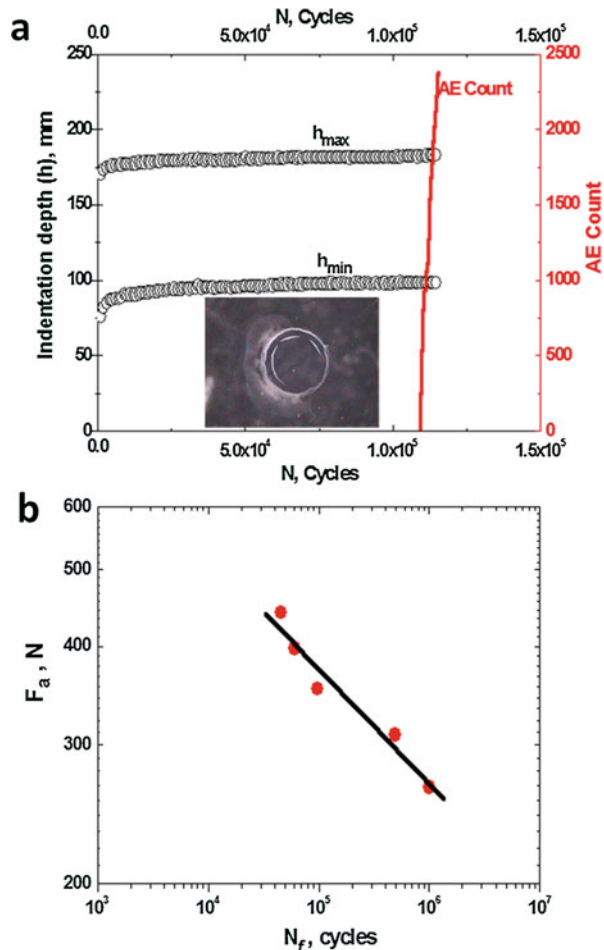
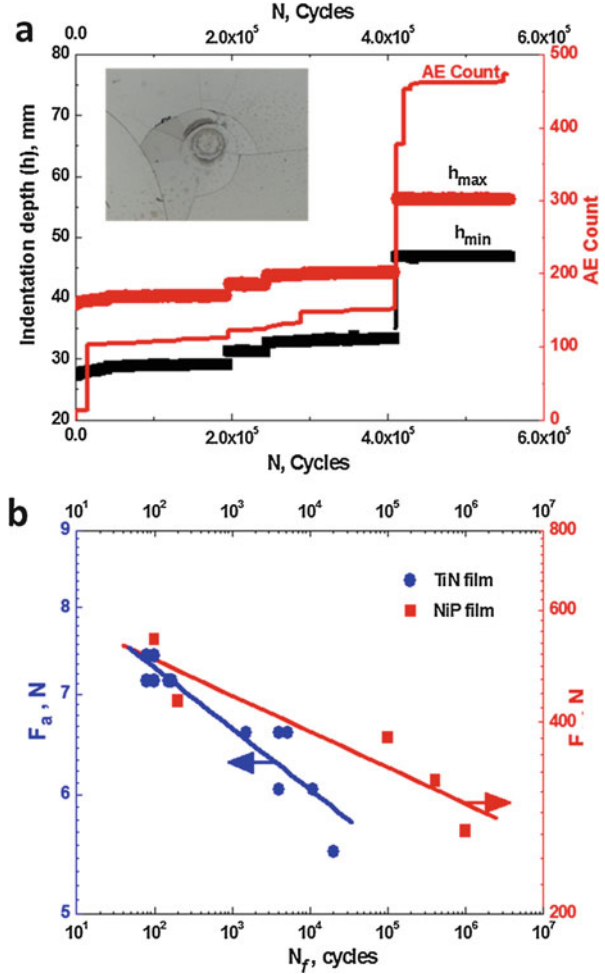


Fig. 13 (a) Experimental evolution of indentation depth and AE count with number of cycle on the NiP thin film/SUS304 steel substrate system. The inset shows the surface topography near the indentation (after failure). (b) Variation of indentation fatigue load amplitude with number of cycles to failure for the Ni-P film/SUS 304 steel substrate system, as well as that of the TiN film/SUS304 substrate system. The maximum indentation fatigue load is 600 N and the load ratio is 0.1. The frequency is 20 Hz Xu et al. 2010 (reprinted with the permission from Taylor & Francis)



relationship is obtained in the log-log plot, which validates the feasibility of using Eq. 25 to describe the indentation fatigue strength of the bulk material.

Next, indentation fatigue on film/substrate systems was also carried out. Figure 13a shows the evolution of indentation depth with number of cycles for the NiP film on a steel substrate, where the maximum indentation force was 600 N. A similar characteristic of the evolution of the indentation depth with that of PVC bulk material is observed: the indentation depth rate remains steady for a number of cycles, and then undergoes a sudden change when the cycle number is about 409,860 cycles. Accordingly, a significant increase of new AE counts (well over 20 per second) was also detected, indicating macrocracking failure in the system. These findings are consistent with the post-observation (inset in Fig. 13a) where the NiP thin film develops many cracks.

Table 1 Comparisons between the fatigue strength exponents obtained from the indentation fatigue test (shaded rows) and uniaxial fatigue test in literature, for several thin film/substrate systems Xu et al. 2010 (reprinted with the permission from Taylor & Francis)

Thin film/substrate	Film thickness (μm)	Fatigue strength exponent n	Error of indentation (present) and tensile (literatures) results
TiN/SUS304	4	-0.041 ± 0.005	–
TiN/AISI 316 L (Puchi-Cabrera et al. 2004)	1.4	-0.037	2.7~24.3%
TiN _{0.75} /AISI 316 L (Berríos et al. 2001)	3	-0.035	2.7~31.4%
NiP/SUS304	200	-0.056 ± 0.013	–
NiP/SAE4340 (D'iaz et al. 2002)	18	-0.084	17.9~48.8%
NiP/AISI 1045 (Contreras et al. 1999)	17	-0.065	6.2~33.8%

With the similar procedures, a series of indentation fatigue tests were also performed with different load amplitudes on the NiP and TiN film/steel substrate systems. Figure 13b shows the relationship between the indentation fatigue load amplitude and the number of cycles to their failure. In both cases, a power-law relationship (i.e., Eq. 25) is validated. Moreover, the indentation fatigue strength exponent, m , can be determined by fitting Fig. 13b, and it equals to -0.041 ± 0.005 and -0.056 ± 0.013 for TiN and NiP thin films on steel substrate, respectively. At the same time, the fatigue strength coefficient F_f can be obtained from the extrapolation and is $8.9 \pm 1\text{N}$ and $660 \pm 120\text{N}$ for TiN and NiP films on steel substrate, respectively, which agrees well the measurements (14.4 N and 1000 N counterparts (Yonezu et al. 2009, 2010)) at the static indentation tests on the same TiN and NiP films on steel substrate, thus suggesting an underlying bridge between the indentation fatigue strength and uniaxial fatigue strength.

Table 1 lists the further comparison of the fatigue stress exponents between the current measurements with those obtained from uniaxial tests. Although the fatigue stress exponents vary with respect to the thickness of films, they are on the same order of magnitude. Therefore, we may conclude that the indentation fatigue strength law, Eq. 25, is suitable for describing the failure behavior with measurement fatigue parameter close to those in conventional uniaxial fatigue test. Note that although the indentation fatigue failure mechanisms of PVC bulk materials, and TiN/SUS304 and NiP/SUS304 film/substrate systems are different, the general potential applicability of Eq. 25 seems to not be affected. This qualitatively echoes the fact that in conventional uniaxial fatigue tests, different failure mechanisms may be involved for different materials, e.g., for brittle and ductile materials, yet Eq. 24 holds for all common mechanisms. Thus, the direct comparison of fatigue strength exponents indicates a quantitative connection between the indentation fatigue test and uniaxial fatigue test.

Concluding Remarks

Indentation fatigue is emerging as an alternative approach to measure the fatigue properties and to probe fatigue-induced deformation and failure mechanism of materials and structures in particular materials and structures with small volumes. This chapter presents, in theory, experiment and computation, extensive evidences of similarity between indentation fatigue depth propagation and fatigue crack growth, and builds a quantitative foundation in mechanics. The recent indentation fatigue experiments performed at the nanoscale on nanomaterials (Cavaliere 2010; Wei et al. 2008) and nanofilms (Bhat 2012) further confirmed the proposed similarity and indentation fatigue laws/equations. Further development of this indentation fatigue technique is envisioned to be integrated with multiple different measurements (e.g., AE) and/or observatory (e.g., TEM, SEM) tools to monitor in situ deformation of materials under cyclic fatigue loadings. At the same time, advanced computational models that can reproduce the experimental results of indentation fatigue will be critically important, and they may require to span several length scales to capture both indentation fatigue phenomena at the macroscale and deformation mechanism at the nanoscale. Moreover, the application of indentation fatigue to measure the cyclic/fatigue properties of broader materials such as soft matter and biological materials and structures will help to extend theories of indentation fatigue mechanics beyond the current focus of engineering metals, and may also offer new opportunities to probe mechanics of soft materials and structures.

References

- M. Abdel-Karim, N. Ohno, Kinematic hardening model suitable for ratchetting with steady-state. *Int. J. Plast.* **16**, 225–240 (2000)
- D.H. Alsem, O.N. Pierron, E.A. Stach, C.L. Muhlstein, R.O. Ritchie, Mechanisms for fatigue of Micron-scale silicon structural films. *Adv. Eng. Mater.* **9**, 15–30 (2007)
- D.H. Alsem, C.L. Muhlstein, E.A. Stach, R.O. Ritchie, Further considerations on the high-cycle fatigue of micron-scale polycrystalline silicon. *Scr. Mater.* **59**, 931–935 (2008)
- S.A.S. Asif, K.J. Wahl, R.J. Colton, Nanoindentation and contact stiffness measurement using force modulation with a capacitive load-displacement transducer. *Rev. Sci. Instrum.* **70**, 2408–2413 (1999)
- X. Baoxing, Y. Zhufeng, C. Xi, Characterization of strain rate sensitivity and activation volume using the indentation relaxation test. *J. Phys. D. Appl. Phys.* **43**, 245401 (2010)
- O.H. Basquin, The exponential law of endurance tests. *Proc. Am. Soc. Test. Mater.* **10**, 625–630 (1910)
- J.A. Berríos, D.G. Teer, E.S. Puchi-Cabrera, Fatigue properties of a 316L stainless steel coated with different TiN_x deposits. *Surf. Coat. Technol.* **148**, 179–190 (2001)
- T. S., Bhat, *Indentation Analysis of Transversely Isotropic Materials*, PhD thesis, Stony Brook University, Stony Brook, 2012
- S. Bhowmick, J.J. Meléndez-Martínez, B.R. Lawn, Bulk silicon is susceptible to fatigue. *Appl. Phys. Lett.* **91**, 201902 (2007)
- L.P. Borrego, J.M. Ferreira, J.M. Pinho da Cruz, J.M. Costa, Evaluation of overload effects on fatigue crack growth and closure. *Eng. Fract. Mech.* **70**, 1379–1397 (2003)

- J.L. Bucaille, S. Stauss, E. Felder, J. Michler, Determination of plastic properties of metals by instrumented indentation using different sharp indenters. *Acta Mater.* **51**, 1663–1678 (2003)
- P. Cavaliere, Cyclic deformation of ultra-fine and nanocrystalline metals through nanoindentation: Similarities with crack propagation. *Procedia Eng.* **2**, 213–222 (2010)
- J.L. Chaboche, D. Nouailhas, Constitutive modeling of ratchetting effects—part I: experimental facts and properties of the classical models. *J. Eng. Mater. Technol.* **111**, 384–392 (1989a)
- J.L. Chaboche, D. Nouailhas, Constitutive modeling of ratchetting effects—part II: possibilities of some additional kinematic rules. *J. Eng. Mater. Technol.* **111**, 409–416 (1989b)
- E.P. Chan, Y. Hu, P.M. Johnson, Z. Suo, C.M. Stafford, Spherical indentation testing of poroelastic relaxations in thin hydrogel layers. *Soft Matter* **8**, 1492–1498 (2012)
- X. Chen, N. Ogasawara, M. Zhao, N. Chiba, On the uniqueness of measuring elastoplastic properties from indentation: The indistinguishable mystical materials. *J. Mech. Phys. Solids* **55**, 1618–1660 (2007)
- R. Chen, Y.C. Lu, F. Yang, G.P. Tandon, G.A. Schoeppner, Impression creep of PMR-15 resin at elevated temperatures. *Polym. Eng. Sci.* **50**, 209–213 (2010)
- Y.T. Cheng, C.M. Cheng, Scaling approach to conical indentation in elastic-plastic solids with work hardening. *J. Appl. Phys.* **84**, 1284–1291 (1998)
- Y.T. Cheng, C.M. Cheng, Scaling relationships in indentation of power-law creep solids using self-similar indenters. *Philos. Mag. Lett.* **811**, 9–16 (2001)
- Y.T. Cheng, C.M. Cheng, Scaling, dimensional analysis, and indentation measurements. *Mater. Sci. Eng. R* **44**, 91–149 (2004)
- S.N.G. Chu, J.C.M. Li, Impression creep; a new creep test. *J. Mater. Sci.* **12**, 2200–2208 (1977)
- S.N.G. Chu, J.C.M. Li, Localized stress relaxation by impression testing. *Mater. Sci. Eng.* **45**, 167–171 (1980a)
- S.N.G. Chu, J.C.M. Li, Delayed retardation of overloading effects in impression fatigue. *J. Eng. Mater. Technol.* **102**, 337–340 (1980b)
- T. Connolly, P.E. McHugh, M. Bruzzi, A review of deformation and fatigue of metals at small size scales. *Fatigue Fract. Eng. Mater. Struct.* **28**, 1119–1152 (2005)
- G. Contreras, C. Fajardo, J.A. Berrío-Aos, A. Pertuz, J. Chitty, H. Hintermann, E.S. Puchi, Fatigue properties of an AISI 1045 steel coated with an electroless Ni-P deposit. *Thin Solid Films* **355–356**, 480–486 (1999)
- J.A. D'íaz, M. Passarelli, J.A. Berrío-Aos, E.S. Puchi-Cabrera, Fatigue behavior of a 4340 steel coated with an electroless Ni-P deposit. *Surf. Coat. Technol.* **149**, 45–56 (2002)
- M. Dao, N. Chollacoop, K.J. VanVliet, T.A. Venkatesh, S. Suresh, Computational modeling of the forward and reverse problems in instrumented sharp indentation. *Acta Mater.* **49**, 3899–3918 (2001)
- R.H. Dauskardt, M.R. James, J.R. Porter, R.O. Ritchie, Cyclic fatigue-crack growth in SiC-whisker-reinforced alumina ceramic composite: Long and small-crack behavior. *J. Am. Ceram. Soc.* **75**, 759–771 (1992)
- P. Delobelle, P. Robinet, L. Bocher, Experimental study and phenomenological modelization of ratchet under uniaxial and biaxial loading on an austenitic stainless steel. *Int. J. Plast.* **11**, 295–330 (1995)
- J.-H. Dirks, E. Parle, D. Taylor, Fatigue of insect cuticle. *J. Exp. Biol.* **216**, 1924–1927 (2013)
- Y. Estrin, A. Vinogradov, Fatigue behaviour of light alloys with ultrafine grain structure produced by severe plastic deformation: an overview. *Int. J. Fatigue* **32**, 898–907 (2010)
- A.C. Fischer-Cripps, *Nanoindentation* (Spring-Verlag, New York, 2000)
- N.A. Fleck, R.A. Smith, Fatigue life prediction of a structural steel under service loading. *Int. J. Fatigue* **6**, 203–210 (1984)
- F. Guibertau, N.P. Padture, H. Cai, B.R. Lawn, Indentation fatigue: a simple cyclic hertzian test for measuring damage accumulation in polycrystalline ceramics. *Philos. Mag. A* **68**, 1003–1016 (1993)
- R.W. Hertzberg, *Deformation and Fracture Mechanics of Engineering Materials* (Wiley, Oxford, 1995)
- R. Hill, *The Mathematical Theory of Plasticity* (Oxford University Press, 1998)

- H.-W. Höppel, H. Mughrabi, A. Vinogradov, *Bulk Nanostructured Materials* (Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2009), pp. 481–500
- Y. Hu, E.P. Chan, J.J. Vlassak, Z. Suo, Poroelastic relaxation indentation of thin layers of gels. *J. Appl. Phys.* **110**, 086103 (2011)
- H.L. Huang, N.J. Ho, The model of crack propagation in polycrystalline copper at various propagating rates. *Mater. Sci. Eng. A* **279**, 254–260 (2000)
- H.L. Huang, N.J. Ho, The observation of dislocation reversal in front of crack tips of polycrystalline copper after reducing maximum load. *Mater. Sci. Eng. A* **345**, 215–222 (2003)
- Y. Jiang, P. Kurath, Characteristics of the Armstrong-Frederick type plasticity models. *Int. J. Plast.* **12**, 387–415 (1996)
- P. Jiang, T. Zhang, Y. Feng, R. Yang, N. Liang, Determination of plastic properties by instrumented spherical indentation: expanding cavity model and similarity solution approach. *J. Mater. Res.* **24**, 1045–1053 (2009)
- K.L. Johnson, *Contact Mechanics* (Cambridge University Press, Cambridge, 1985)
- P. Kaszynski, E. Ghorbel, D. Marquis, An experimental study of ratchetting during indentation of 316L stainless steel. *J. Eng. Mater. Technol.* **120**, 218–223 (1998)
- R. Kumar, A. Kumar, S. Kumar, Delay effects in fatigue crack propagation. *Int. J. Press. Vessel. Pip.* **67**, 1–5 (1996)
- H. Lan, T.A. Venkatesh, Determination of the elastic and plastic properties of materials through instrumented indentation with reduced sensitivity. *Acta Mater.* **55**, 2025–2041 (2007)
- B. Lawn, R. Wilshaw, Indentation fracture: principles and applications. *J. Mater. Sci.* **10**, 1049–1081 (1975)
- C. Lee, X. Wei, J.W. Kysar, J. Hone, Measurement of the elastic properties and intrinsic strength of monolayer graphene. *Science* **321**, 385–388 (2008)
- J.H. Lee, T. Kim, H. Lee, A study on robust indentation techniques to evaluate elastic–plastic properties of metals. *Int. J. Solids Struct.* **47**, 647–664 (2010)
- J.C.M. Li, Impression creep and other localized tests. *Mater. Sci. Eng. A* **322**, 23–42 (2002)
- X.D. Li, B. Bhushan, A review of nanoindentation continuous stiffness measurement technique and its applications. *Mater. Charact.* **48**, 11–36 (2002)
- J.C.M. Li, S.N.G. Chu, Impression fatigue. *Scr. Metall.* **13**, 1021–1026 (1979)
- W.B. Li, J.L. Henshall, R.M. Hooper, K.E. Easterling, The mechanisms of indentation creep. *Acta Metall. Mater.* **39**, 3099–3110 (1991)
- X. Li, H. Gao, C.J. Murphy, K.K. Caswell, Nanoindentation of silver nanowires. *Nano Lett.* **3**, 1495–1498 (2003)
- P. Li, Q. Liao, S. Yang, X. Bai, Y. Huang, X. Yan, Z. Zhang, S. Liu, P. Lin, Z. Kang, Y. Zhang, In situ transmission electron microscopy investigation on fatigue behavior of single ZnO wires under high-cycle strain. *Nano Lett.* **14**, 480–485 (2014)
- S.Y. Liu, I.W. Chen, Fatigue of yttria-stabilized zirconia - I. Fatigue damage, fracture origins and lifetime prediction. *J. Am. Ceram. Soc.* **74**, 1197–1205 (1991)
- S.J. Lloyd, A. Castellero, F. Giuliani, Y. Long, K.K. McLaughlin, J.M. Molina-Aldareguia, N.A. Stelmashenko, L.J. Vandeperre, W.J. Clegg, Observations of nanoindents via cross-sectional transmission electron microscopy: a survey of deformation mechanisms. *Proc. R. Soc. A* **461**, 2521–2543 (2005)
- J.L. Loubet, W.C. Oliver, B.N. Lucas, Measurement of the loss tangent of low-density polyethylene with a nanoindentation technique. *J. Mater. Res.* **15**, 1195–1198 (2000)
- J. Luo, K. Dahmen, P.K. Liaw, Y. Shi, Low-cycle fatigue of metallic glass nanowires. *Acta Mater.* **87**, 225–232 (2015)
- D. Mclean, *Mechanical Properties of Metals* (Wiley Press, New Jersey, 1965)
- P. Miranzo, J.S. Moya, Elastic/plastic indentation in ceramics: a fracture toughness determination method. *Ceram. Int.* **10**, 147–152 (1984)
- J.C. Moosbrugger, D.J. Morrison, Nonlinear kinematic hardening rule parameters – direct determination from completely reversed proportional cycling. *Int. J. Plast.* **13**, 633–668 (1997)
- K.A. Nibur, D.F. Bahr, Identifying slip systems around indentations in FCC metals. *Scr. Mater.* **49**, 1055–1060 (2003)

- J.D. Nowak, K.A. Rzepiejewska-Malyska, R.C. Major, O.L. Warren, J. Michler, In-situ nanoindentation in the SEM. *Mater. Today* **12**(Supplement 1), 44–45 (2010)
- N. Ohno, J. Wang, On modelling of kinematic hardening for ratcheting behaviour. *Nucl. Eng. Des.* **153**, 205–212 (1995)
- W.C. Oliver, G.M. Pharr, Measurement of hardness and elastic modulus by instrumented indentation: advances in understanding and refinements to methodology. *J. Mater. Res.* **19**, 3–20 (2004)
- E.S. Puchi-Cabrera, F. Mat'inez, I. Herrera, J.A. Berr'ios, S. Dixit, D. Bhat, On the fatigue behavior of an AISI 316L stainless steel coated with a PVD TiN deposit. *Surf. Coat. Technol.* **182**, 276–286 (2004)
- G.D. Quinn, R.C. Bradt, On the Vickers indentation fracture toughness test. *J. Am. Ceram. Soc.* **90**, 673–680 (2007)
- Y. Rao, R.J. Farris, Fatigue and creep of high-performance fibers: deformation mechanics and failure criteria. *Int. J. Fatigue* **30**, 793–799 (2008)
- R.O. Ritchie, Influence of microstructure on near-threshold fatigue crack propagation in ultra-high strength steel. *Met. Sci.* **11**, 368–381 (1977)
- R.O. Ritchie, Mechanisms of fatigue-crack propagation in ductile and brittle solids. *Int. J. Fract.* **100**, 55–83 (1999)
- K. Sadananda, A.K. Vasudevan, Fatigue crack growth mechanisms in steels. *Int. J. Fatigue* **25**, 899–914 (2003)
- S. Sakaguchi, N. Murayama, Y. Kodama, & F. Wakai, in *Fracture Mechanics of Ceramics: Fracture Fundamentals, High-Temperature Deformation, Damage, and Design*, eds. by R. C. Bradt, D. P. H. Hasselman, D. Munz, M. Sakai, V. Ya Shevchenko (Springer US, 1992), pp. 509–521
- I.N. Sneddon, The relationship between load and penetration in the axisymmetric Boussinesq problem for a punch of arbitrary profile. *Int. J. Eng. Sci.* **3**, 47–57 (1965)
- D.S. Stone, J.E. Jakes, J. Puthoff, A.A. Elmustafa, Analysis of indentation creep. *J. Mater. Res.* **25**, 611–621 (2010)
- B. Storakers, P.L. Larsson, On Brinell and Boussinesq indentation of creeping solids. *J. Mech. Phys. Solids* **42**, 307–332 (1994)
- S. Suresh, *Fatigue of materials* (Cambridge University Press, Cambridge, 1998)
- B. Taljat, G.M. Pharr, Development of pile-up during spherical indentation of elastic-plastic solids. *Int. J. Solids Struct.* **41**, 3891–3904 (2004)
- Y. Tang, A. Yonezu, N. Ogasawara, N. Chiba, X. Chen, On radial crack and half-penny crack induced by Vickers indentation. *Proc. R. Soc. A: Math. Phys. Eng. Sci.* **464**, 2967–2984 (2008)
- J. Tang, J. Li, J.J. Vlassak, Z. Suo, Fatigue fracture of hydrogels. *Extreme Mech. Lett.* **10**, 24–31 (2017)
- S.H. Teoh, Fatigue of biomaterials: a review. *Int. J. Fatigue* **22**, 825–837 (2000)
- V. Tvergaard, Overload effects in fatigue crack growth by crack-tip blunting. *Int. J. Fatigue* **27**, 1389–1397 (2005)
- O.L. Warren, Z. Shan, S.A.S. Asif, E.A. Stach, J.W. Morris Jr., A.M. Minor, In situ nanoindentation in the TEM. *Mater. Today* **10**, 59–60 (2007)
- P.J. Wei, Y.C. Wang, J.F. Lin, Retardation of cyclic indentation creep exhibited in metal alloys. *J. Mater. Res.* **23**, 2650–2656 (2008)
- Z. Xia, W.A. Curtin, B.W. Sheldon, A new method to evaluate the fracture toughness of thin films. *Acta Mater.* **52**, 3507–3517 (2004)
- B. Xu, X. Chen, Determining engineering stress–strain curve directly from the load–depth curve of spherical indentation test. *J. Mater. Res.* **25**, 2297–2307 (2010)
- B. Xu, Z. Yue, Study of the ratcheting by the indentation fatigue method with a flat cylindrical indenter: part I. Experimental study. *J. Mater. Res.* **21**(7), 1793–1797 (2006)
- B. Xu, Z. Yue, Study of the ratcheting by the indentation fatigue method with a flat cylindrical indenter: part II. Finite element simulation. *J. Mater. Res.* **22**, 186–192 (2007)
- B. Xu, B. Zhao, Z. Yue, Investigation of residual stress by the indentation method with the flat cylindrical indenter. *J. Mater. Eng. Perform.* **15**, 299–305 (2006)
- B. Xu, Z. Yue, J. Wang, Indentation fatigue behaviour of polycrystalline copper. *Mech. Mater.* **39**(12), 1066–1080 (2007a)

- B. Xu, X. Wang, Z. Yue, Indentation behavior of polycrystalline copper under fatigue peak overloading. *J. Mater. Res.* **22**, 1585–1592 (2007b)
- B.X. Xu, X.M. Wang, B. Zhao, Z.F. Yue, Study of crystallographic creep parameters of nickel-based single crystal superalloys by indentation method. *Mater. Sci. Eng. A* **478**, 187–194 (2008a)
- B.X. Xu, X.M. Wang, Z.F. Yue, Determination of the internal stress and dislocation velocity stress exponent with indentation stress relaxation test. *J. Mater. Res.* **23**, 2486–2490 (2008b)
- B. Xu, Z. Yue, X. Chen, An indentation fatigue depth propagation law. *Scr. Mater.* **60**(10), 854–857 (2009)
- B. Xu, Z. Yue, X. Chen, Numerical investigation of indentation fatigue on copper polycrystalline. *J. Mater. Res.* **24**(3), 1007–1015 (2009)
- B.X. Xu, A. Yonezu, X. Chen, An indentation fatigue strength law. *Philos. Mag. Lett.* **90**(5), 313–322 (2010)
- F. Yang, J.C.M. Li, Impression creep by an annular punch. *Mech. Mater.* **21**, 89–97 (1995)
- F. Yang, J.C.M. Li, Impression test – a review. *Mater. Sci. Eng. R. Rep* **74**, 233–253 (2013)
- W. Yang, S. Mao, J. Yang, T. Shang, H. Song, J. Mabon, W. Swiech, J.R. Vance, Z. Yue, S.J. Dillon, H. Xu, B. Xu, Large-deformation and high-strength amorphous porous carbon nanospheres. *Sci. Rep.* **6**, 24187 (2016)
- A. Yonezu, B. Xu, X. Chen, Indentation induced lateral crack in ceramics with surface hardening. *Mater. Sci. Eng. A* **507**, 226–235 (2009)
- A. Yonezu, B. Xu, X. Chen, An experimental methodology for characterizing fracture of hard thin films under cyclic contact loading. *Thin Solid Films* **8**, 2082–2089 (2010)
- N. Zaaferani, D. Raabe, F. Roters, S. Zaefferer, On the origin of deformation-induced rotation patterns below nanoindentations. *Acta Mater.* **56**, 31–42 (2008)
- Y. Zhu, C. Ke, H.D. Espinosa, Experimental techniques for the mechanical characterization of one-dimensional nanostructures. *Exp. Mech.* **47**, 7 (2007)



Crack Initiation and Propagation in Laminated Composite Materials

14

Jun Xu and Yanting Zheng

Contents

Introduction	434
Constitutive Relations	435
Incremental Form of Constitutive Relations When the Material Is Not Damaged	437
Incremental Form of Constitutive Relations When the Material Is Damaged	440
Verification of Constitutive Relations	440
Numerical Simulation	442
PVB Windshield FEA Modeling Based on Constitutive Relations of PVB	
Laminated Glass	442
Numerical Simulation Based on the Extended Finite Element Method	447
Experiment Analysis	458
Experiment of Quasi-Static Loading	458
Experiments of Dynamic Out-of-Plane Loading	462
Summary	492
References	493

Abstract

PVB laminated glass is a kind of typical laminated composite material and its crack characteristics are of great interest to vehicle manufacturers, safety engineers, and accident investigators. Because crack morphology on laminated

J. Xu (✉)

Department of Automotive Engineering, School of Transportation Science and Engineering,
Beihang University, Beijing, China

Advanced Vehicle Research Center (AVRC), Beihang University, Beijing, China
e-mail: junxu@buaa.edu.cn

Y. Zheng

China Automotive Technology and Research Center, Tianjin, China
e-mail: yantingzheng@buaa.edu.cn

windshield contains important information on energy mitigation, pedestrian protection, and accident reconstruction. In this chapter, we investigated the propagation characteristics for both radial and circular cracks in PVB laminated glasses by theoretical constitutive equations analysis, numerical simulation, experiments, and tests of impact. A damage-modified nonlinear viscoelastic constitutive relations model of PVB laminated glass were developed and implemented into FEA software to simulate the pedestrian head impact with vehicle windshield. Results showed that shear stress, compressive stress, and tensile stress were main causes of plastic deformation, radial cracks, and circumferential cracks for the laminated glass subject to impactor. In addition, the extended finite element method (XFEM) was adopted to study the multiple crack propagation in brittle plates. The effects of various impact conditions and sensitivity to initial flaw were discussed. For experiment analysis, crack branching was investigated and an explicit expression describing the crack velocity and number of crack branching is proposed under quasi-static Split Hopkinson Pressure Bar (SHPB) compression experiments. And the radial crack propagation behavior of PVB laminated glass subjected to dynamic out of - plane loading was investigated. The steady-state cracking speed of PVB laminated glass was lower pure glass, and it increased with higher impactor speed and mass. The supported glass layer would always initiate before the loaded layer and the final morphologies of radial cracks on both sides are completely overlapped. Two different mechanisms of crack propagation on different glass layers explained the phenomenon above. Then further parametric dynamic experiments study on two dominant factors, i.e., impact velocity and PVB thickness are investigated: Firstly, a semiphysical model describing the relationship between the maximum cracking velocity and influential factors was established; Then the Weibull statistical model was suggested considering various factors to describe the macroscopic crack pattern in this chapter; Finally, the relation between radial crack velocity and crack numbers on the backing glass layer and the relation between the crack length and the capability of energy absorption on the impacted glass layer were proposed.

Keywords

PVB Laminated windshield · Crack Propagation and Initiation · Theoretical Constitutive Relations · Numerical Simulation of Crack Propagation · Extended Finite Element Method · Experiments Analysis of Crack Propagation · Quasi-Static Loading · Dynamic Out-of-Plane Loading

Introduction

As an typical laminated composite material, the standard PVB laminated glass comprises of a PVB interlayer sandwiched by two mono soda-lime and it has extensive applications in architecture, automobile industry, as well as the structural parts during the past few decades where they usually play an important role in human protection and structural integrity due to its excellent energy-absorbing and

fragment-holding ability. However, mechanical behavior especially for the cracks of PVB laminated glass subject to loading has not been fully explored yet due to the complicity caused by the composite material, i.e., polymer material sandwiched by two pieces of brittle material. As a result, it becomes a pressing need to investigate the mechanical loading information contained in the macroscopic crack initiation and propagation which enables improved design or prevention for laminated PVB glass as a structural part.

In this chapter, there will be three methods mainly applied to investigate PVB laminated glass damage mechanism and crack growth: theoretical constitutive equations describing the failure criteria of PVB laminated glass under load, numerical simulations of crack initiation and growing on PVB laminated glass, experiments and tests of impact between impactor and laminated glass plate. For theoretical investigation, a damage-modified nonlinear viscoelastic constitutive equation is developed based on the updated Lagrangian approach. For numerical simulation investigation, a FEA model with proper material parameters verified by a classical example is established to describing the impact between pedestrian head and PVB laminated windshield. In addition, the extended finite element method (XFEM) is employed to study the low-speed impact-induced cracking in brittle plates in the present chapter. Upon head impact, the interaction between the stress field and the initiation and propagation of the radial and circumferential cracks are computed using XFEM. The effects of various impact conditions and sensitivity to initial flaw are discussed. The experiment analysis of PVB laminated glass is mainly based on quasi-static loading and dynamic impact. We carried out quasi-static Split Hopkinson Bar (SHPB) experiments to study the crack velocity and number of crack branching. For the dynamic out-of-plane impact tests carried out to study the crack propagation behavior of PVB laminated glass subjected to dynamic impact by using high-speed photography, the time histories of the averaged radial crack tip position, propagation velocity, and acceleration are recorded and the parametric study of impactor speed and mass, impact velocity, PVB thickness for PVB laminated glass are investigated. Then two different mechanisms of crack propagation on different glass layers are explained. Besides, a semiphysical model describing the relationship between the maximum cracking velocity and influential factors is established. In addition, the Weibull statistical model is suggested considering various factors to describe the macroscopic crack pattern thus providing a theoretical evidence for engineering practice. Finally, the relation between radial crack velocity and crack numbers on the backing glass layer and the relation between the crack length and the capability of energy absorption on the impacted glass layer are investigated.

Constitutive Relations

The constitutive relation of laminated composite material can well reflect the mechanical behavior of that. In order to study the crack in laminated composite material, the constitutive relation should be established firstly.

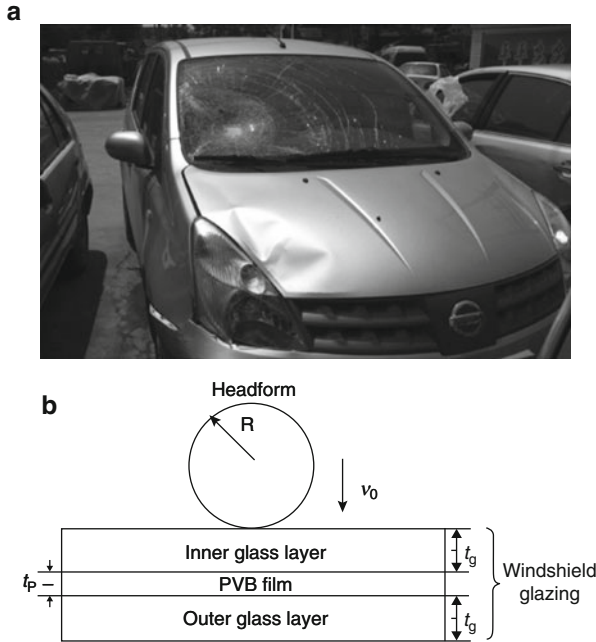


Fig. 1 (a) PVB windshield crack morphology in real-world pedestrian-vehicle accident. (b) Schematic illustration of a sphere pedestrian headform impacting on a windshield glazing

PVB windshield used in automotive industry are a kind of laminated composite material which comprise of a PVB interlayer sandwiched by two momo soda-lime glass sheet. Thus, pedestrian head impacts on the vehicle PVB windshield can be regarded as a scenario where a composite plate is subjected to a concentrated force by a spherical indenter on its top surface shown in Fig. 1. The composite plate consists of $n = 2$ individual brittle soda-lime glass layers of thickness t_g and modulus E_g , sandwiched with a PVB interlayer of thickness t_p and modulus E_p .

By studying the constitutive relations of the PVB windshield, the crack in PVB windshield can be analyzed thoroughly. Lili et al. (Wang et al. 2003; Lili et al. 1991) study the mechanical behavior of polymer with performing dynamic impact and quasi-static tensile experiments. According to the experimental results, the behavior of the polymer can be described as the following constitutive relation:

$$\sigma = \sigma_e + \sigma_{t_1} + \sigma_{t_2} \tag{1}$$

σ_t stand for the tensile stress.

$$\sigma_e = \sigma_m \left[1 - \exp \left(- \sum_{t=1}^n \frac{(m\varepsilon)^t}{t} \right) \right] \tag{2}$$

$$\sigma_{t_1} = E_1 \int_0^t \dot{\varepsilon}(\tau) \exp\left(-\frac{t-\tau}{\theta_1}\right) d\tau \quad (3)$$

$$\sigma_{t_2} = E_2 \int_0^t \dot{\varepsilon}(\tau) \exp\left(-\frac{t-\tau}{\theta_2}\right) d\tau \quad (4)$$

where m , τ , E_1 , E_2 are a nondimensional relative initial elastic modulus, relaxation time as defined in Maxwell body, elastic constants of first Maxwell body and elastic constants of second Maxwell body, respectively. θ_1 , θ_2 , are constants as well, standing for the relaxation times of first and second Maxwell body. Where σ_e stands for the stress with the damage material. σ_{t_1} , σ_{t_2} standing for tensile stress of first and second Maxwell body. D is a “damage” variable.

Nevertheless, the above constitutive relations are only valid when strain is less than 7% (Lili et al. 1991), and it cannot express the relation completely. To further utilize the above relations, Fenghua et al. (Zhou et al. 1992) suggested the following equation to describe the constitutive relations:

$$\sigma = (1 - D) \sigma_e + \sigma_{t_1} + \sigma_{t_2} \quad (5)$$

where D is a “damage” variable and is defined as:

$$D = \begin{cases} 0 & \varepsilon \leq \varepsilon_{th} \\ \dot{D}_0 \dot{\varepsilon}^{\delta-1} (\varepsilon - \varepsilon_{th}) & \varepsilon > \varepsilon_{th} \end{cases} \quad (6)$$

And ε_{th} here regards as strain threshold of damage evolution. \dot{D}_0 , δ both are material parameters determined by experiments. Then the dynamic damage criterion is put forward:

$$D_f = \dot{D}_0 \dot{\varepsilon}^{\delta-1} (\varepsilon - \varepsilon_{th}) \quad (7)$$

Incremental Form of Constitutive Relations When the Material Is Not Damaged

When employing the constitutive relations, we should consider some conditions like whether the strain is larger than the threshold strain. Thus we need an equivalent strain to determine whether it is or not. As following is the equivalent:

$$\varepsilon_{eq} = \frac{\sqrt{2}}{3} \sqrt{(\varepsilon_1 - \varepsilon_2)^2 + (\varepsilon_2 - \varepsilon_3)^2 + (\varepsilon_3 - \varepsilon_1)^2 + \frac{3}{2}(\gamma_{12}^2 + \gamma_{23}^2 + \gamma_{31}^2)} \quad (8)$$

where subscripts “1,” “2,” and “3” denote the directions of principal stresses.

On the basis of updated Lagrangian virtual work method, stress-strain relations of material that undergoes large deformation with the ignorance of physical nonlinear properties can be regarded as (Fischer and Washizu 1982):

$$\{S_{ij}\} = \int_0^t E(t - \tau) [A] \frac{\partial \{e_{kl}\}}{\partial \tau} d\tau \tag{9}$$

$\{S_{ij}\}$, $\{e_{kl}\}$ are Kirchhoff stress tensors and Green strain tensors, respectively.

$$[A] = \frac{1-\nu}{(1+\nu)(1-2\nu)} \cdot \begin{bmatrix} 1 & \frac{\nu}{(1-\nu)} & \frac{\nu}{(1-\nu)} & 0 & 0 & 0 \\ \frac{\nu}{(1-\nu)} & 1 & \frac{\nu}{(1-\nu)} & 0 & 0 & 0 \\ \frac{\nu}{(1-\nu)} & \frac{\nu}{(1-\nu)} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1-2\nu}{2(1-\nu)} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1-2\nu}{2(1-\nu)} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1-2\nu}{2(1-\nu)} \end{bmatrix} \tag{10}$$

Similarly, as PVB windshield is a three-dimensional structure, Equations from 1 to 6 can be rewritten as (Swanson and Christensen 2015; Wang et al. 2007; Shen et al. 1987):

$$S_{ij} = S_{ij,e} + S_{ij,t_1} + S_{ij,t_2} \tag{11}$$

where

$$S_{ij,e} = \sigma_m \left\{ 1 - \exp \left(-mA E_{kl} - \sum_{t=1}^n \frac{(mE_{ij})^t}{t} \right) \right\} \tag{12}$$

$$S_{ij,t_1} = E_1 \int_0^t A \frac{\partial E_{kl}}{\partial \tau} \exp \left(-\frac{t - \tau}{\theta_1} \right) d\tau \tag{13}$$

$$S_{ij,t_2} = E_2 \int_0^t A \frac{\partial E_{kl}}{\partial \tau} \exp \left(-\frac{t - \tau}{\theta_2} \right) d\tau \tag{14}$$

On the basis of Boltzmann superposition (Feng 1995), incremental form of constitutive relations can be obtained. Also, implementation in commercial FEA software needs the incremental form of the stress tensor of Eqs. 12, 13, and 14 to be rewritten. Thus, it is a must for us to reach the incremental form first. According to proofs of Chain Rule, Eq. 12 can be rewritten as follows:

$$dS_{ij,e} = \sigma_m \exp \left(-mA E_{kl} - \sum_{t=2}^n \frac{(mE_{ij})^t}{t} \right) \cdot \left[mA \frac{\partial E_{kl}}{\partial \tau} + \sum_{t=2}^n \frac{(mE_{ij})^{t-1}}{t} \frac{\partial E_{kl}}{\partial \tau} \right] \cdot \Delta t \tag{15}$$

Considering Δt is short enough yields (Feng 1995):

$$\frac{\partial E_{kl}}{\partial \tau} = \frac{\partial E_{kl}}{\partial t} = \frac{\partial E_{kl}}{\Delta \tau} \quad (16)$$

Thus Eq. 13 becomes:

$$\Delta S_{ij,e} = \sigma_m \exp\left(-mA E_{kl} - \sum_{t=2}^n \frac{(mE_{ij})^t}{t}\right) \cdot \left[mA \Delta E_{kl} + \sum_{t=2}^n \frac{(mE_{ij})^{t-1}}{t} \Delta E_{kl} \right] \quad (17)$$

Similarly, viscoelastic stress tensors for both low strain rate and high strain rate at t are:

$$\Delta S_{ij,t_1} = E_1 \theta_1 A \frac{1 - \exp\left(-\frac{\Delta t}{\theta_1}\right)}{\Delta t} \Delta E_{kl} - \left(1 - \exp\left(-\frac{\Delta t}{\theta_1}\right)\right)^t S_{ij,t_1} \quad (18)$$

$$\Delta S_{ij,t_2} = E_2 \theta_2 A \frac{1 - \exp\left(-\frac{\Delta t}{\theta_2}\right)}{\Delta t} \Delta E_{kl} - \left(1 - \exp\left(-\frac{\Delta t}{\theta_2}\right)\right)^t S_{ij,t_2} \quad (19)$$

Therefore, the incremental form of stress tensor during one unit time step Δt can be obtained:

$$\begin{aligned} \Delta S_{ij,n+1} &= \Delta S_{ij,e} + \Delta S_{ij,t_1} + \Delta S_{ij,t_2} = \sigma_m \exp\left(-mA E_{kl} - \sum_{t=2}^n \frac{(mE_{ij})^t}{t}\right) \\ &\cdot \left[mA \Delta E_{kl} + \sum_{t=2}^n \frac{(mE_{ij})^{t-1}}{t} \Delta E_{kl} \right] + E_1 \theta_1 A \frac{1 - \exp\left(-\frac{\Delta t}{\theta_1}\right)}{\Delta t} \Delta E_{kl} \\ &- \left(1 - \exp\left(-\frac{\Delta t}{\theta_1}\right)\right)^t S_{ij,t_1} + E_2 \theta_2 A \frac{1 - \exp\left(-\frac{\Delta t}{\theta_2}\right)}{\Delta t} \Delta E_{kl} \\ &- \left(1 - \exp\left(-\frac{\Delta t}{\theta_2}\right)\right)^t S_{ij,t_2} \end{aligned} \quad (20)$$

In the following way, the stress tensor can be calculated:

$$S_{ij,n+1} = S_{ij,n} + \Delta S_{ij,n+1} \quad \varepsilon < \varepsilon_{th} \quad (21)$$

where $t_{n+1} = t_n + \Delta t$.

Incremental Form of Constitutive Relations When the Material Is Damaged

When the material is not damaged, the incremental form of constitutive relations can be described using function relation above. However, when the strain is larger than the threshold strain, material turns into “damage” phase. As mentioned above, stress tensor in “damage” phase at time t_{n+1} can be concluded as:

$$\{S_{ij}\}_{n+1} = \{S_{ij}\}_n + \{\Delta S_{ij}\}_{n+1} \quad (22)$$

and Eq. 21 can be rewritten as:

$$S_{ij,n+1} = S_{ij,n} + \Delta S_{ij,n+1} \quad (23)$$

Therefore,

$$\begin{aligned} S_{ij,n+1} &= S_{ij,n} + \Delta S_{ij,n+1} \\ &= \frac{1}{1-D_{n+1}} S_{ij,n+1}^d + 1^d - \frac{1}{1-D_n} S_{ij,n}^d \\ &= \frac{1}{1-D_{n+1}} \Delta S_{ij,n+1}^d + \frac{\Delta D_{n+1} S_{ij,n}^d}{(1-D_{n+1})(1-D_n)} \end{aligned} \quad (24)$$

where $\Delta D_{n+1} = D_{n+1} - D_n$ is the increase of damage variable.

Rewriting Eq. 23 yields:

$$\begin{aligned} \Delta S_{ij,n+1}^d &= (1 - D_{n+1}) \Delta S_{ij,n+1} - \frac{\Delta D_{n+1} S_{ij,n}^d}{(1 - D_n)} \\ &= (1 - D_{n+1}) \Delta S_{ij,n+1} - \Delta D_{n+1} \Delta S_{ij,n} \end{aligned} \quad (25)$$

Similarly, the stress tensor under damage can be calculated as:

$$S_{ij,n+1}^d = S_{ij,n}^d + \Delta S_{ij,n+1}^d \quad \varepsilon \geq \varepsilon_{th} \quad (26)$$

Verification of Constitutive Relations

The constitutive relation of PVB windshield are gotten preliminary and the accuracy of that should be studied next. In order to verify the above constitutive relations, a classical example is chosen. A thin plate with infinite length and $2b = 12$ mm in width is under the uniform load $q(t) = 1$ kN·m. There is a small crack in center with length $2a = 1$ mm, see Fig. 2. Material properties are listed in Table 1.

Fig. 2 Thin plate with center crack under uniform load

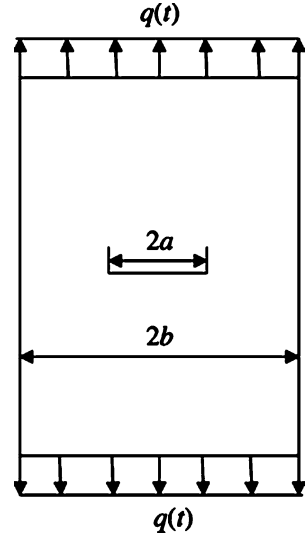


Table 1 Material parameters for plate

σ_m /MPa	n	m	E_1 /MPa	θ_1/s	E_2 /MPa	$\theta_2/\mu s$	ε_{th}	δ	D_0	D_f
1,100	4	19.9	949	13.8	3,981	67.4	0.055	1.22	0.863	0.0797

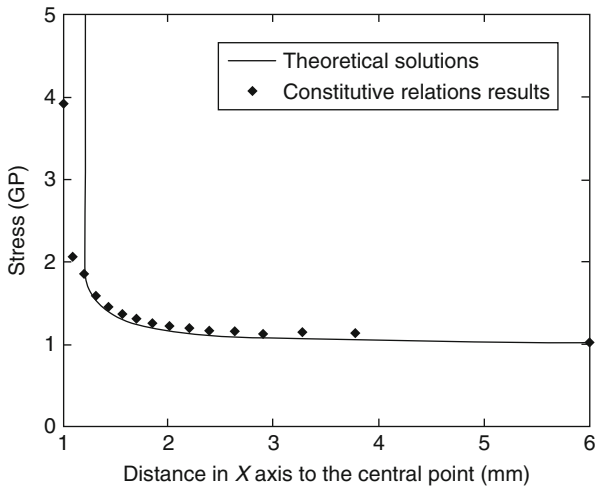


Fig. 3 A comparison between theoretical solutions and FEA results

By implementing the above constitutive relations into FEA software numerical and adding the command of failure of material, results can be obtained. Comparing numerical results with theoretical ones, we can see that the constitutive relations results can coincide with the theoretical solutions well, shown in Fig. 3.

Numerical Simulation

PVB Windshield FEA Modeling Based on Constitutive Relations of PVB Laminated Glass

By verifying the constitutive relations above with a classical example, the relations have a sufficient accuracy. Then the FEA modeling of PVB windshield are introduced in detail following.

Contact Analysis

A nonconforming two-body contact model of a PVB windshield loaded by a headform (sphere) is illustrated in Fig. 4. The thickness of PVB windshield is 8.76 mm, the length of it is 350 mm, and the thickness of interlayer PVB is 0.76 mm. Shell elements were chosen to mesh the two bodies in contact. Hourglass control and automatic surface- to-surface contact with dynamic friction coefficient 0.1 were employed in LS-Dyna. Initial nodal gaps between the two contact surfaces were prescribed. The contact elements were not activated until the penetration of the sphere into the composite glass occurred for the sake of saving computation time.

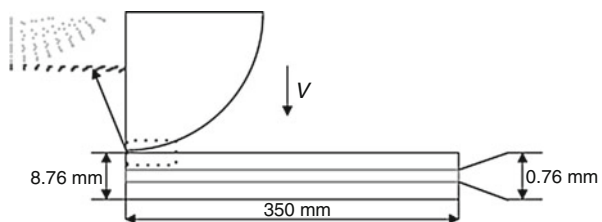
An extra command that allows the failure of material is added in the input file of LS-Dyna. Then the above-mentioned constitutive relations with damage variable could be fully employed in FEA. The accuracy of the contact FEA model is firstly examined by classical Hertzian pressure calculation method. Illustration of classical Hertzian theory is shown in Fig. 5 (Daphalapurkar et al. 2007). The δ and a_0 stand for mutual approach between two bodies and their contact radius.

Deviations between FEA and theoretical results are all within 1.5%. Comparison shown in Fig. 6 demonstrates the sufficient accuracy of the FEA model to stimulate the impact between pedestrian head and PVB windshield.

Material Model

The PVB laminated glass is modeled according to the following principles: if the glass does not fail, the composite is treated as a shell; otherwise, the interlayer PVB film acts as a membrane. A Belytschko–Tsay shell element for the glass material and a membrane element for the interlayer are employed (Zhou et al. 1992; Fischer

Fig. 4 A simplified physical model of impact between head and windshield



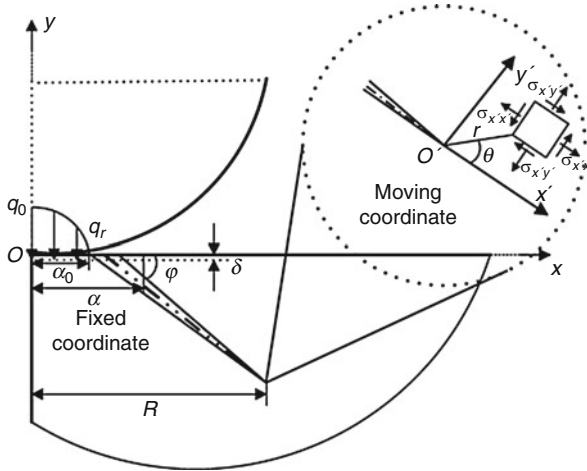


Fig. 5 A simplified crack-model based on Hertzian pressure with moving coordinate at the crack tip is shown

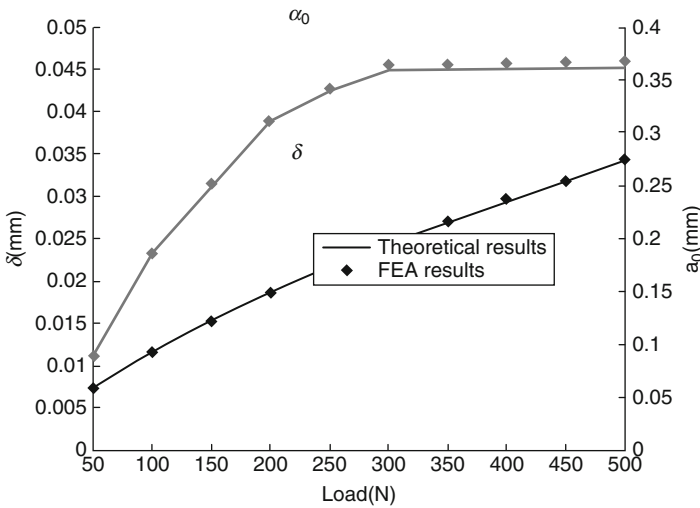


Fig. 6 Comparison between FEA results and theoretical results

and Washizu 1982). Both types of elements are fully integrated in simulations so that no hourglass modes should be expected (Sharon and Fineberg 1999).

Glass and PVB film both can be treated as isotropic materials according to their physical and material properties. Therefore, we assume that each layer of composite windshield is isotropic with respect to its material symmetry lines and obeys Hooke's law. It is assumed that, due to the influence of dynamic loading, there is no sliding between two contact layers. It then becomes reasonable to consider the

Table 2 Parameters used in both constitutive relations and finite element analysis

Components	Parameters and values
Headform	$E = 65 \text{ GPa}, \rho = 1412 \text{ kg/m}^3, \nu = 0.22$
Glass	$E = 74 \text{ GPa}, \rho = 2500 \text{ kg/m}^3, \nu = 0.25, t_g = 2 \text{ mm}$
PVB film	$k = 20 \text{ GPa}, \rho = 1100 \text{ kg/m}^3, t_p = 0.76 \text{ mm}$
Windshield dimension	Panel dimensions (a × b): 1,320 × 630 mm, with a small curvature in x and z direction in coordinates
Other configuration	Initial velocity: $v_0 = 10 \text{ m/s}$ with no gravity field

composite windshield as integrity. As a result, an equivalent elastic modulus and Poisson's ratio should be introduced as follows (Dwivedi and Espinosa 2003).

$$\bar{E} = \frac{2t_g E_g + t_p E_p}{2t_g + t_p} \quad (27)$$

$$\bar{\nu} = \frac{2t_g \nu_g + t_p \nu_p}{2t_g + t_p} \quad (28)$$

where \bar{E} and $\bar{\nu}$ are equivalent Young's modulus and equivalent Poisson's ratio of windshield glazing.

Material parameters are listed in Table 2.

Results and Discussions

Internal Stress Analysis

The cracks in PVB windshield are demonstrations of internal and external stresses. Thus, the FEA model above is used to analyze the internal stresses. Only through investigating the internal stresses can we know the cracks on windshields better.

During the dynamic crack evolution shown in FEA analysis, stages of damage evolution could be sketchily drawn as Fig. 7. Three kinds of cracks are: circumferential cracks, web crack, and plastic crack. The circumferential cracks are mainly caused by pressure stress while the web cracks are mainly caused by tensor stress. Plastic cracks are the last stage that the windshield deforms plastically. In Fig. 8, three kinds of internal stresses, tensile stress σ_t , compressive stress σ_c , and shear stress σ_s , are plotted versus impact time. We can find out that the value of stresses sorted from large to small is σ_s , σ_c and σ_t . Lawn and Wilshaw (1975) stated that in addition to the Hertzian cone crack and the median vent crack, there is also an inelastic deformation zone where shear and hydrostatic compression is greatest. Quite similarly, it is obvious to see in Fig. 8 that shear stress is the greatest among three stresses, causing the windshield to deform plastically.

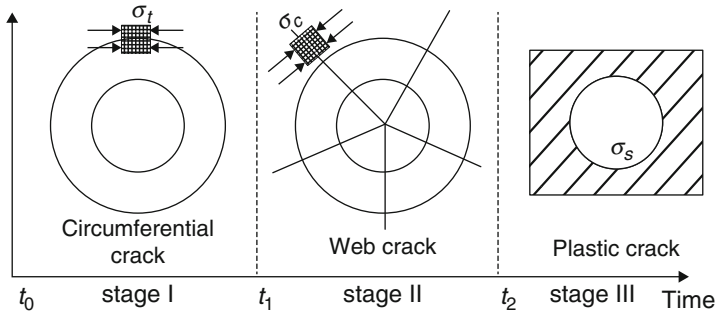


Fig. 7 Diagram of different crack-growing stage

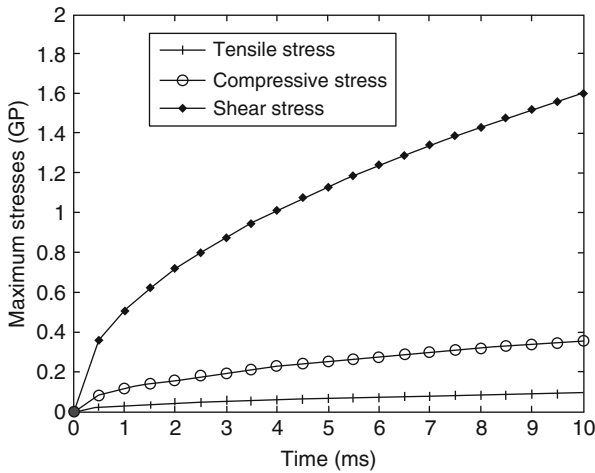


Fig. 8 Three kinds of maximum internal stress in PVB windshield, tensile stress, compressive stress, and shear stress curves versus impact time

Internal Stress Variation Among Different Point on Windshield Plate

Internal stresses of different points on windshield plate are far different and crack information can prove it. The cracks become weaker as the radius grows larger for circumferential cracks, as do the radial cracks. In Fig. 9, maximum stresses appear at the very point that impacted by pedestrian head. Stresses decrease as the distances from the central point increase. The decreasing rates of three stresses are nearly the same according to the computational results.

Effects of Poisson’s Ratio on Crack Angle

Poisson’s ratio plays an important role of cone crack angle. A parametric study is performed to manifest the effects on crack angle. Poisson’s ratio is ranged from 0.2 to 0.4, with a step-width of 0.05. Table 3 shows the results of FEA analysis

Fig. 9 Three kinds of maximum internal stress in PVB windshield, tensile stress, compressive stress, and shear stress curves versus the displacement from the central point

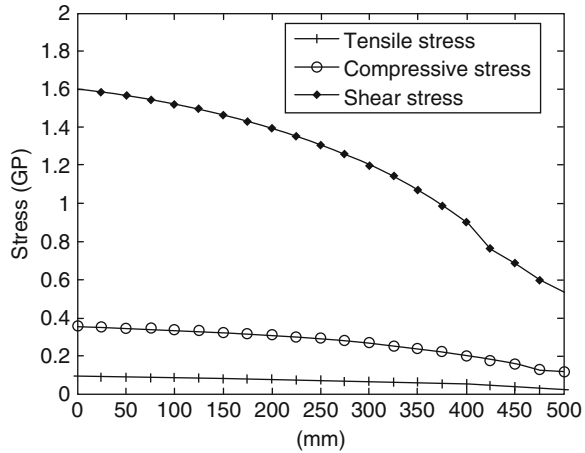


Table 3 Comparison results of two models on Poisson's ratio effect to crack angle

Poisson's ratio	0.20	0.25	0.30	0.35	0.40
SIF analysis	12.4	15.3	19.6	24.5	27.7
FEA analysis	11.8	15.1	18.9	23.1	26.2
Differences (%)	5.20	1.32	3.70	6.10	5.80

compared to that of stress intensity factor analysis (SIF). Maximum difference between the results is below 7%. With the increase in Poisson's ratio, the conical crack angle also increases though the increase pattern is still unknown.

Effects of Impact Velocity

Vehicle speed is a critical factor in pedestrian-vehicle traffic accident. Extends and degrees of injuries largely depend on the impact velocity between pedestrian and vehicle. Therefore, studies of cracks on windshields under different head impact velocity would provide the essential foundations in pedestrian protection. According to the census data in NTADTU, in most cases, vehicles impact pedestrians at the speed of about 40–70 km/h. Thus, we vary the impact velocity of the head from 0 to 20 m/s and make impact velocity V_0 as an independent variable. On the other hand, we choose the maximum radius of circumferential cracks R_c , the maximum length of radial cracks L_r , and the maximum radius of plastically deformed area R_p as dependent variables. Three dependent variables are illustrated and defined in Fig. 10.

In Fig. 11, the impact velocity has very great influence to three chosen variables. At the impact velocity of 20 m/s, R_c can reach up to about 642 mm, much larger than $R_c \approx 89$ mm under the impact velocity of 5 m/s. Vehicle velocity plays an extremely important role in both pedestrian head injury and damage of windshield. In addition, it is reasonable to infer the extend of injuries of pedestrian head and impacting velocity of vehicle roughly through R_c , L_r , and R_p .



Fig. 10 Effect of initial impact velocity of the stress

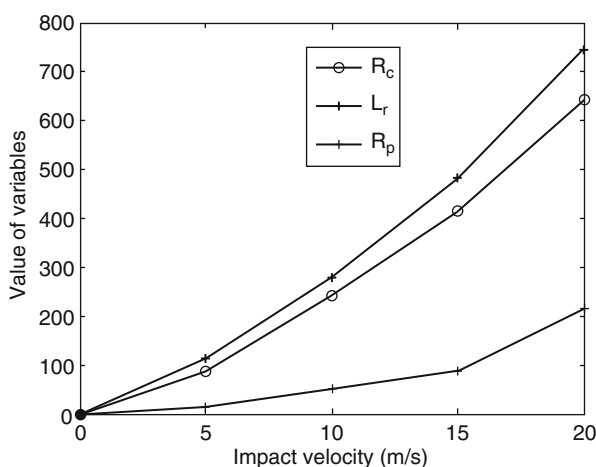


Fig. 11 Effect of initial impact velocity to cracks

Numerical Simulation Based on the Extended Finite Element Method

After the head impacts on the windshield material, cracks including radial crack and circumferential crack appear in the glass material. The fracture characteristics of the windshield are widely recognized as one of the most important factors in automotive crashworthiness. The crack profiles (e.g., length, pattern, etc.) on the windshield material contain critical information for impact speed (which is extremely useful for accident reconstruction), vehicle crashworthiness, and insights for improving pedestrian and passenger protection. This motivates us to study the characteristics of low-speed head impact-induced crack propagation on the windshield glass material. In the past studies of impact on windshield glass, due to the intrinsic complexity

of numerical analyses of cracking, researchers in automotive engineering often avoid explicit simulation of crack propagation. The following study aims to study the crack propagation characteristics (including both radial and circumferential cracks) when a model windshield undergoes low-speed impact and the numerical investigation of multiple crack propagation is based on the extended finite element method (XFEM) in this study.

Fundamentals of XFEM

XFEM incorporates a discontinuous displacement field across the crack facing away from the crack tip, in the form of:

$$u^h = \sum_{i=1}^n N_i(x) \left(u_i + b_i H(x) + \sum_{l=1}^4 \sum_{j=1}^2 c_{il} F_l^j(x) \right) \quad (29)$$

where n is the number of nodes in the mesh, $N_i(x)$ is the shape function of node i , u_i are the classical DOFs of node i . b_i and c_{il} are the DOFs associated with the Heaviside “jump” function $H(x)$, with value 1 above the crack and below the crack. The crack-tip function $F_l^j(x)$ is:

$$F_l(r, \theta) = \left(\sqrt{r} \sin\left(\frac{\theta}{2}\right), \sqrt{r} \cos\left(\frac{\theta}{2}\right), \sqrt{r} \sin\left(\frac{\theta}{2}\right) \sin \theta, \sqrt{r} \cos\left(\frac{\theta}{2}\right) \sin \theta \right) \quad (30)$$

where (r, θ) are the local polar coordinates at the crack tip.

Figure 12 illustrates the enriched nodes near a crack tip, where $H(x)$ is discontinuous across the crack surface. Under general mixed mode-loadings, the asymptotic near-tip hoop and shear stress components are

$$\begin{cases} \sigma_{\theta\theta} = \frac{1}{4\sqrt{2\pi r}} (k_1 (3 \cos(\theta/2) + \cos(3\theta/2)) + k_{II} (-3 \sin(\theta/2) - 3 \sin(3\theta/2))) \\ \sigma_{\theta\theta} = \frac{1}{4\sqrt{2\pi r}} (k_1 (\sin(\theta/2) + \sin(3\theta/2)) + k_{II} (\cos(\theta/2) - 3 \cos(3\theta/2))) \\ \sigma_{\theta\theta} = \frac{1}{4\sqrt{2\pi r}} (k_1 (5 \cos(\theta/2) - \cos(3\theta/2)) + k_{II} (5 \sin(\theta/2) - 3 \sin(3\theta/2))) \end{cases} \quad (31)$$

For fracture propagation, we use the popular criterion of maximum local hoop stress, where the crack propagation direction is determined from Eq. 31 as

$$\theta_c = \arccos \left(\frac{3k_{II}^2 + \sqrt{K_I^4 + 8k_I^2 k_{II}^2}}{K_I^2 + 9k_{II}^2} \right) \quad (32)$$

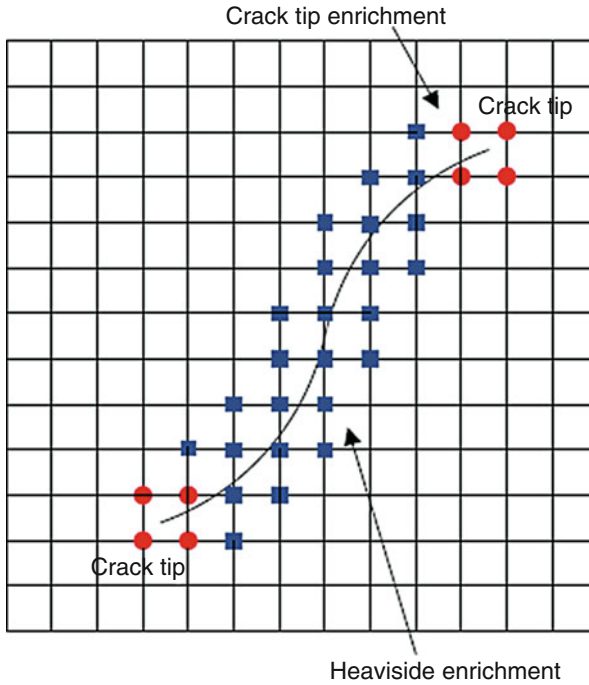


Fig. 12 Schematic of uniform mesh surrounding a crack used in XFEM simulation. *Blue square nodes* stand for the Heaviside function enrichment and *red circular nodes* stand for the crack tip enrichment. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article)

Here, θ_c is measured with respect to a local polar coordinate system with its origin at the crack tip and aligned with the direction of the existing crack. Numerical simulations based on the above XFEM algorithm is carried out using commercial code ABAQUS.

Model and Methods

Model Setup

A featureless spherical headform is used to model the human head; The model head has a radius $R = 90$ mm and mass 4.5 kg. Note that effective head weights are changeable in different accident cases because human body may also get involved due to inertia, causing the effective impact mass larger than the mass of the pure head, especially during higher speed impact. Thus, the effective head mass M is also varied from 4.5 kg up to 90 kg by varying the density (from 1,500 to 30,000 kg/m³) and keeping volume the same in this study. According to 151 pedestrian–vehicle accident cases from the National Traffic Accident Database of Tsinghua University (NTADTU), most of the impact speeds of the vehicle in accident fall into the range of 40–80 km/h, i.e. 11.1–22.2 m/s, and that is the range of low-speed impact we

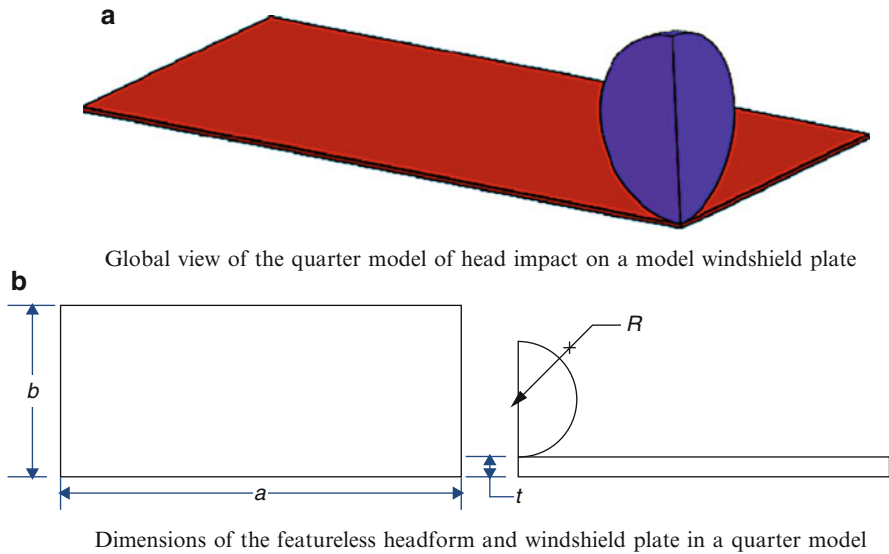


Fig. 13 Computational model setup. (a) Global view of the quarter model of head impact on a model windshield plate. (b) Dimensions of the featureless headform and windshield plate in a quarter model

consider. We focus on normal impact and assume that impact occurs at the center of the plate, we neglect the small windshield curvature and focus on a flat rectangular plate with dimensions $a \times b = 700 \times 300$ mm and thickness $t = 4.76$ mm as Fig. 13 shows.

In terms of the material properties, typical material parameters of glass are assigned to the plate, with density $\rho = 2,500$ kg/m³, Young’s modulus to the plate, Poisson’s ratio $\nu = 0.22$, mode I energy release rate $G_I = 10$ J/m², mode II and III energy release rate $G_{II} = G_{III} = 50$ J/m². The properties are assumed to be rate independent. To take advantage of symmetry, only one quarter of the model is needed for simulation. The boundary condition of the windshield is assumed to be fixed. The plate and head are modeled with eight-node linear brick elements. The contact between the head and plate obeys the Coulomb’s friction law with coefficient $f = 0.1$ (since according to our accident scene investigation, a small slip exists between head and windshield).

Computational Method

Suppose there is no crack or damage, we first compute the deformation response with response to impact. For a given effective head mass M and impact speed v , the maximum displacement D is computed and presented in a dimensionless form in Fig. 14. For the range of impact speed and head mass considered in this study, D/t can be fitted as a dimensionless function of v/v_0 and M/M_0 , as

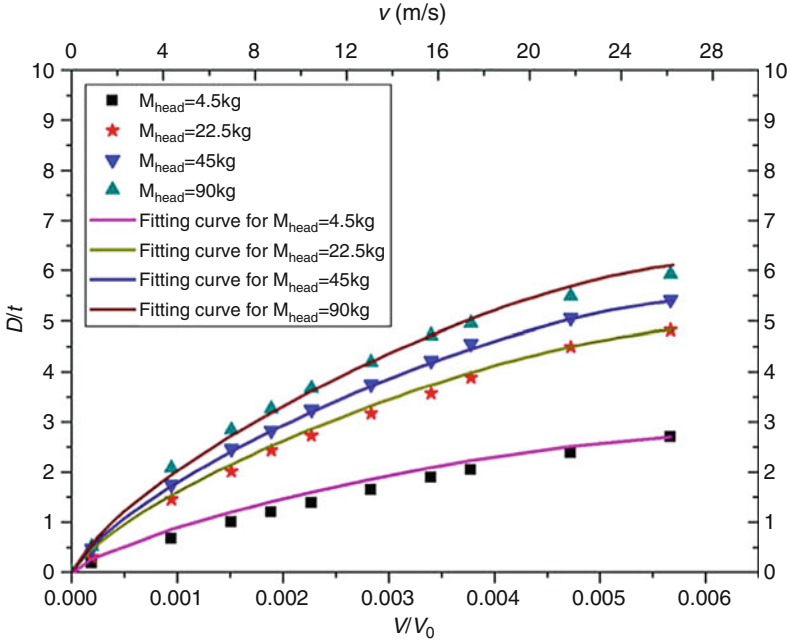


Fig. 14 Relationship between the initial impact velocity and displacement under different head masses; *lines* denote curve fitting, Eq. 33

$$D/t = \left(135.3(v/v_0)^{0.7115} - 0.0305^{518.5(v/v_0)} \right) (M/M_0)^{0.2669} \quad (33)$$

where $v_0 = \sqrt{E/\rho}$ and $M_0 = abt \rho$.

Since the impact speed is relatively low (much smaller than the wave speed), our simulation shows that the dynamic effect is relatively minor for the range of parameters investigated in this chapter. That is, for a given head mass, the computed stress field from dynamic impact (with incident speed v) is very close to that of quasi-static indentation (with maximum indentation depth D/t) as long as they induce the same deformation curvature of the plate (i.e., when v and D/t satisfy Eq. 33); see an example in Fig. 15 for the stress field. This enables us to simplify the dynamic impact problem into a quasi-static indentation one, where the normalized indentation depth D/t becomes the governing variable to indicate the effective “impact condition.”

Our previous study showed that for windshield fracture, circumferential cracks always grow after the radial crack has developed, except near the vicinity of impact where extensive damage can be found (Xu et al. 2009). This is consistent with Fig. 15 that the hoop stress field is more prominent than radial stress. Thus, the crack simulation can be decoupled into two steps. First, we embed a number of

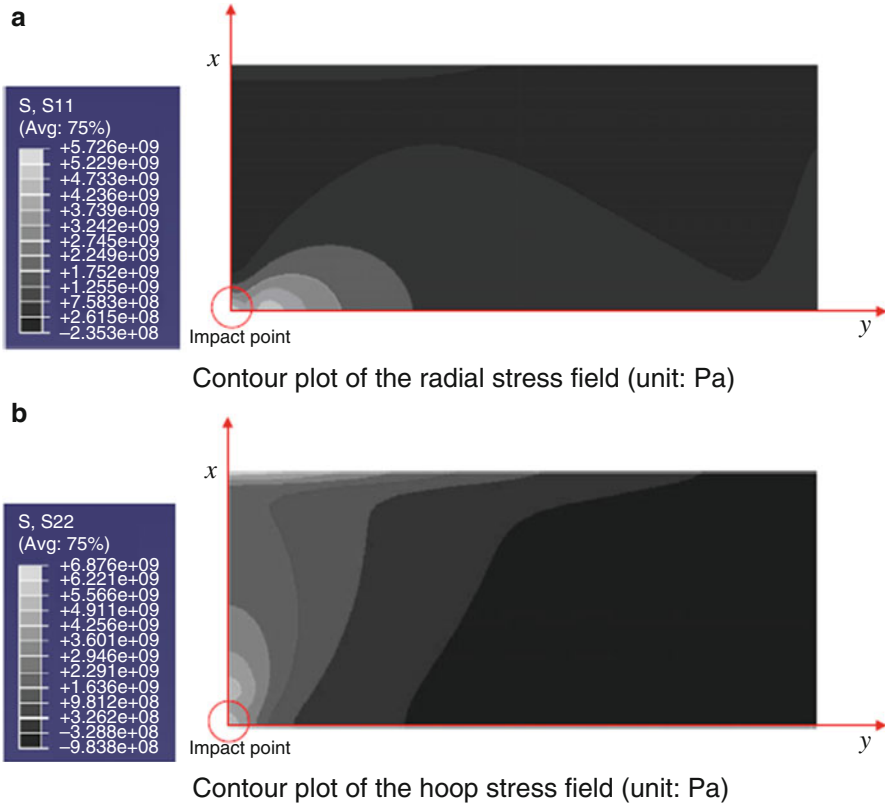


Fig. 15 Comparison between radial and hoop stress field resulted purely from impact deformation; to relieve the prominent stress, cracks need to form and radial crack is preferred to propagate before the circumferential crack owing to the more significant hoop stress field. **(a)** Contour plot of the radial stress field (unit: Pa). **(b)** Contour plot of the hoop stress field (unit: Pa)

initial flaws at the center of the plate (impact location); this is consistent with the fact that extensive damage is initiated in the contact zone and that facilitates subsequent crack initiation and propagation. The size of the initial flaw is $a_0 = 3$ mm with variable directions. When these initial flaws are very small, only the radial flaws will grow to become radial cracks upon impact, so as to release strain energy. For a given indentation depth D , the radial crack grows gradually until a certain length; with the increase of D , the final length of radial crack also increases. In the second step, based on the radially cracked specimen, new circumferential flaws with size $a_0 = 3$ mm are introduced (the sensitivity of the flaw geometry and distribution is discussed below), and without further increasing the load, with the residual stress and strain fields inherited from the first step, the development of circumferential cracks is simulated by XFEM.

Rack Propagation Characteristics

Radial Crack Propagation Characteristics

Overall Crack Feature

In Fig. 16, for a typical displacement of $D/t = 0.62$, the evolved radial crack is shown along with the hoop stress field. Note that with the fully developed crack, this stress field is already much relaxed compared to its counterpart without cracking (see for example the difference between Figs. 15 and 16); moreover, the current hoop stress field is specified for the global coordinate, whereas during crack propagation, it is the local hoop stress that determines the crack trajectory. When the local maximum hoop stress exceeds a critical threshold, referred to as the damage stress σ_d , the crack grows incrementally toward such a local principal direction.

In the beginning stage of impact, since the growth of radial crack can release more energy than the circumferential crack (due to the more prominent hoop stress field, Fig. 15), radial crack grows with the increase of the indentation depth D/t . It is noticeable that the crack growth direction keeps changing slightly with growth direction 22° approximately. This is mainly caused by the rectangular shape of the windshield plate, which makes the overall strain in the y -direction larger than that in the x -direction. When the crack grows further, the boundary constraint changes and the crack direction approaches about 45° .

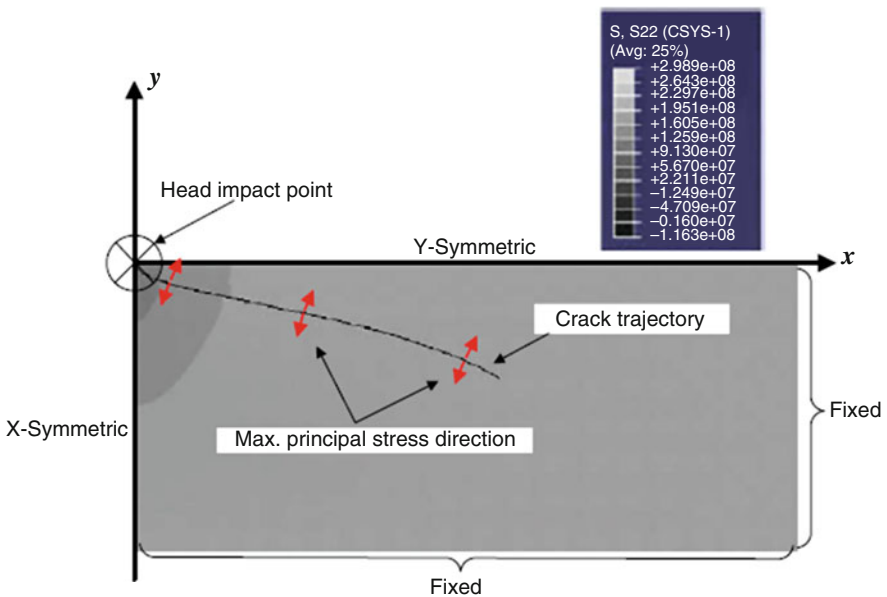


Fig. 16 View of crack trajectory and the hoop stress field after the radial crack is fully developed at $D/t = 0.62$

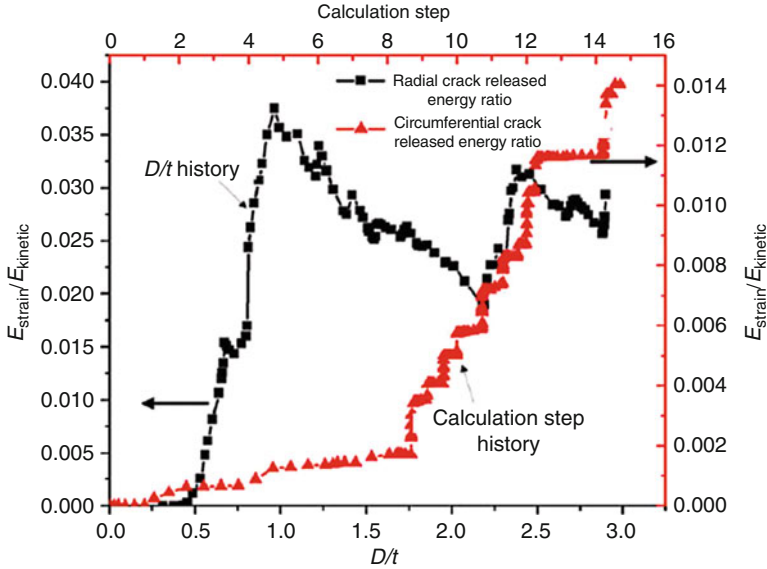


Fig. 17 The normalized released strain energy during the propagation of radial and circumferential cracks. For radial crack, the released strain energy is a function of D/t (or equivalently the impact speed through Eq. 33); it is normalized by the incident kinetic energy with $M = 4.5$ kg. The circumferential crack propagates at $D/t = 2.89$ and it is normalized by the corresponding incident kinetic energy

During its growing process, the crack seeks to release as much energy as possible through mode I fracture. The released strain energy E_{strain} is computed as the difference between the strain energies of two specimens, one cracked and one without crack (at the same D/t). In Fig. 17, E_{strain} is normalized by the initial impact kinetic energy of the head E_{kinetic} (with $M = 4.5$ kg, the impact speed can be deduced from Eq. 33). It is also illustrated that the released strain energy is relatively small compared with the overall system energy, and it increases as D/t enlarges (although when it is normalized by the incident kinetic energy, which also increases as D/t gets larger, the dimensionless number $E_{\text{strain}}/E_{\text{kinetic}}$ fluctuates somewhat).

Sensitivity of Initial Flaw and Fracture Criterion

The damage stresses σ_d of different soda-lime glass specimens are different due to its intrinsic stochastic flaws, usually from 10 to 60 MPa. Under quasi-static indentation, the effect of σ_d on crack evolution is shown in Fig. 18. The results indicate that as the critical stress for crack initiation σ_d becomes larger, the crack angle is smaller (the boundary constraint is more prominent) and the crack is shorter. In other words, if the impact condition is known, one can roughly estimate the damage stress threshold value according to the crack path and length. In what follows, $\sigma_d = 10$ MPa is employed.

Fig. 18 Effect of different damage stresses on the radial crack trajectory, for $D/t = 0.62$

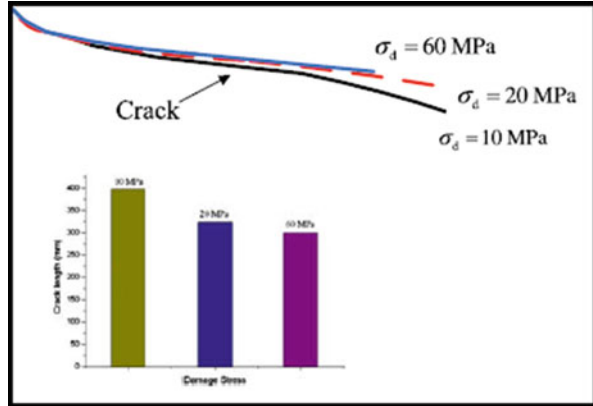
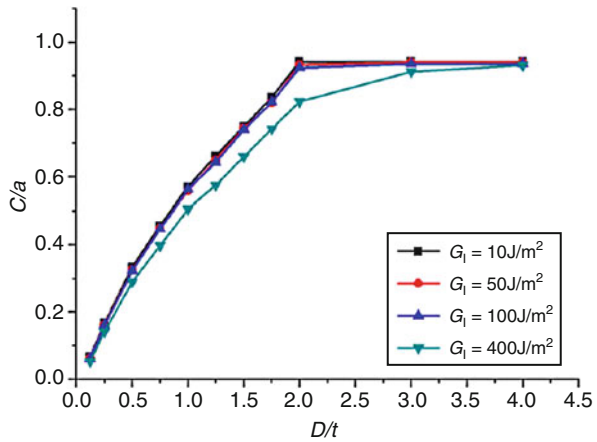


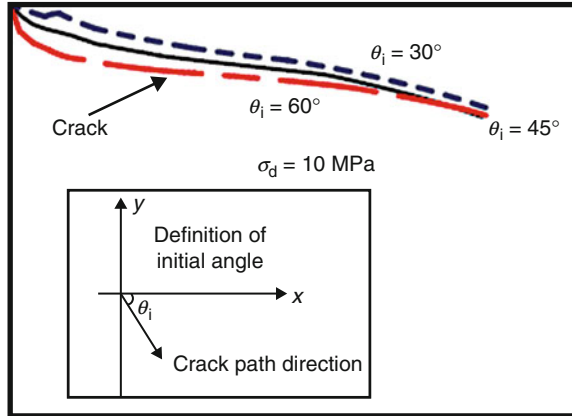
Fig. 19 The effect of model I energy release rate effect on the radial crack length



The model I energy release rate is another very important parameter governing the brittleness of the material. Consider typical material candidates for windshield such as soda-lime glass and PMMA, the former is a brittle material whereas the later has certain degree of ductility, their mode I energy release rates are about $G_I = 10 \text{ J/m}^2$ and $G_I = 400 \text{ J/m}^2$, respectively. Without loss of generality, we normalized the crack length C with plate size a . In Fig. 19, the effect on the resulting length of crack is examined by varying G_I from 10 to 400 J/m^2 . Numerical study demonstrates that if the G_I value is within the brittle material domain, there is little effect on the crack length; otherwise, if the windshield is made by the more ductile material, the crack length is shorter under the same plate deformation. In what follows, $G_I = 10 \text{ J/m}^2$ is employed.

The characteristics of initial flaw have some minor effects on crack propagation. We first keep the initial flaw length to be 3 mm and let the flaw angle θ_i to be 30° , 45° , and 60° with respect to the x-axis, and the final crack pattern is shown in Fig. 20. It is observed that θ_i does have some influence on the initial crack trajectory;

Fig. 20 The effect of the initial flaw angle θ_i on the radial crack pattern, for $D/t = 0.62$



however, eventually, all crack paths converge with similar lengths, indicating both the robustness of the current model. The length of the initial flaw has negligible effect on both the crack propagation pattern and crack length (as long as the initial flaw is short). When the initial flaw length is taken to be 0.5 mm, 1 mm, 3 mm, and 5 mm, the resulting radial crack lengths in these four scenarios are exactly the same; if the initial flaw is 10 mm, the crack propagation pattern is slightly different. Therefore, the 3 mm initial flaw used in the present study is validated and it is a relatively insensitive parameter in the present XFEM simulation.

Radial Crack Length

The length of the radial crack can be represented as a function of the quasi-indentation depth D/t , which is in turn related to the impact speed v for a given effective head mass (see Eq. 34). Figure 19 also shows the normalized crack lengths, i.e., C/a versus the normalized head displacement D/t . Initially, the crack length is almost proportional to head displacement; however, after a certain length owing to the fixed boundary, the crack length almost remains a constant. The relationship between C/a and D/t is nonlinear and may be fitted as

$$C/a = -0.4046 + 0.9231 (D/t) - 0.05074(D/t)^3 + 0.03803 \exp(D/t) - 0.4243 \sqrt{(D/t)} \ln(D/t) \tag{34}$$

By combining Eqs. 33 and 34, a new method can be derived to estimate the impact speed directly based on the knowledge of radial crack length (assuming the effective head mass is known) using Newton–Raphson method, thus useful for traffic accident reconstruction. This is illustrated in the next section.

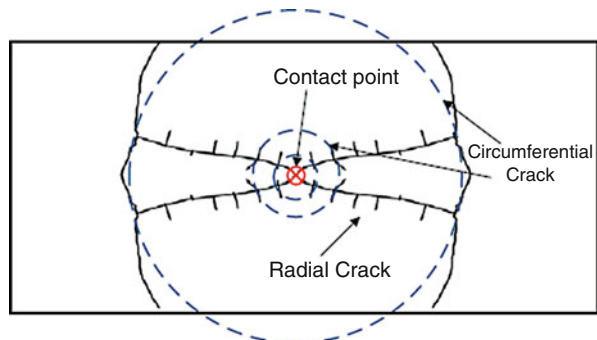
Equation 34 shows that if the head mass is fixed, the crack length increases nonlinearly with impact speed. In order to estimate the influence of the effective head mass, we can combine Figs. 14 and 19, along with Eqs. 33 and 34. For instance, from Fig. 14 (or Eq. 33), if the impact speed is fixed at 5 m/s, for different effective

head masses, the variation of crack length (as a function of the corresponding D/t) can be obtained from either Fig. 19 or Eq. 34. Compared with the crack length of $M = 22.5$ kg, $M = 45$ kg, and $M = 90$ kg its 193.27%, 241.35%, and 294.23% for $M = 4.5$ kg, respectively. That is to say, the effective head mass has a strong influence over the crack length, indicating that during a low-speed impact accident, the shape of vehicle front end and pedestrian–vehicle contact point is critical in impact speed estimation since these factors will largely determine the pedestrian revolution angle. In addition, the influence of M becomes relatively smaller when M is large, implying that the influence of the effective mass starts to decay if the effective head mass is much larger than the windshield mass.

Circumferential Crack Propagation Characteristics

On the basis of the radially cracked geometry, in the second step, we introduce circumferential flaws of length $a_0 = 3$ mm along the radial crack, with a spacing of 50 mm (the spacing is found to be a relatively minor factor on the final circumferential crack pattern). The stress field calculated from the end of first step is transferred into the second step. Since the first step has relaxed most of the hoop stresses, the remaining excessive radial stress needs to be relaxed via the development of circumferential crack, and the most efficient way of releasing the radial stress is to develop the circumferential crack with maximum radius, i.e., on the same order of C . This is verified from Fig. 21: Cracks also grow near the contact zone because of the high stress. This is consistent with the observation that intensive cracks are identified within the contact area while the crack density is lower outside, and the largest circumferential crack almost bounds the radial cracks. The strain energy released by circumferential crack is given in Fig. 17, where the total released energy is less than 50% than that of radial crack. This evidence strongly supports the argument that crack is prone to grow in a radial way rather than a circumferential way, since more energy can be released. From Fig. 17, we can see that less than about 5% of the total kinetic impact energy is consumed by both radial cracks and circumferential cracks and thus there is enough room for improving the automotive windshield for enhanced energy mitigation efficiency and pedestrian protection (for example, using transparent energy absorption materials).

Fig. 21 Final crack pattern (including both radial and circumferential cracks) on the model windshield plate subject to head impact with $D/t = 0.62$



Experiment Analysis

The mechanical behavior of PVB laminated glass is quite complex, especially under dynamic loadings due to the combination of rate-dependent effect of PVB and brittle-polymer composite structure. Therefore, experimental data are very important for the proper suggestion of constitution model and FE model to further investigation mechanical behavior of PVB laminated glass under both static and dynamic loadings.

Typically, one can categorize experiments for studying mechanical response of such a composite material into three regimes in terms of different strain rates: quasi-static, low impact speed, and high impact speed (Stout et al. 1999).

Experiment of Quasi-Static Loading

Because of the difficulties in conducting a quasi-static tension experiment with a slippery surface and the windshield undergoes compression subject to human head impact, we adopted a compression experiment rather than the bending test which the shear effect will influence the mechanical property of the specimen to a large extent.

Experimental Setup

PVB laminated glass circular flake is introduced to be the testing specimen, with the dimension of $u = 10$ mm, where 0.76 mm-thick PVB interlayer is sandwiched by two pieces of 2 mm-thick glass, shown in Fig. 22. Note that the glass, as a typical rate-independent brittle material, has a density of $\rho = 2,500$ kg/m³, Young's modulus of $E = 70$ GPa, and Poisson's ratio of $m = 0.22$ (Forquin and Hild 2010). It is worthy to point out that PVB laminated glass is made by putting two glass sheets into an autoclave and compressed with 0.76 mm-thick of PVB interlayer under specific cycle of pressure (10 bar) and heat (120 °C). Thus, we can regard there is no slide between three layers (Wang et al. 2007). To make the experiment more engineering applicable, the thickness of specimen used here has the same thickness as that in most passenger cars. In experiment, specimen is put in a cylinder container which has a steel pin with a diameter $u = 38$ mm inside for compression and the experiments are carried out on Material Testing System (MTS). The load is applied perpendicularly through the pin to the sample. A clip-gage type extensometer is installed between the gaps to measure the displacement of the pin as a function of load during the experiment, shown in Fig. 23. 1 mm/min, 0.1 mm/min, and 0.01 mm/min are three loading rates, corresponding to the strain rate on the upper and lower sides of the specimen of $\dot{\epsilon} \approx 1 \times 10^{-3} s^{-1}$, $\dot{\epsilon} \approx 1 \times 10^{-4} s^{-1}$, $\dot{\epsilon} \approx 1 \times 10^{-5} s^{-1}$ chosen to measure the quasi-static mechanical property of PVB laminated glass. During tests, unloading threshold value is set to be 80% of the maximum loading value and thus complete load–displacement history is recorded. The temperature and relative humidity during the experiments remain in the range of 25–26 °C and 40–45%, respectively.

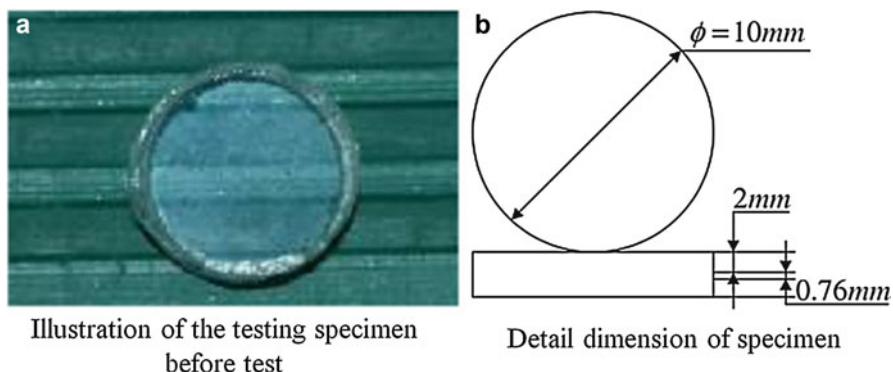


Fig. 22 Illustrations and dimensions of specimen used in both quasi-static and dynamic impact experiments. (a) Illustration of the testing specimen before test. (b) Detail dimension of specimen

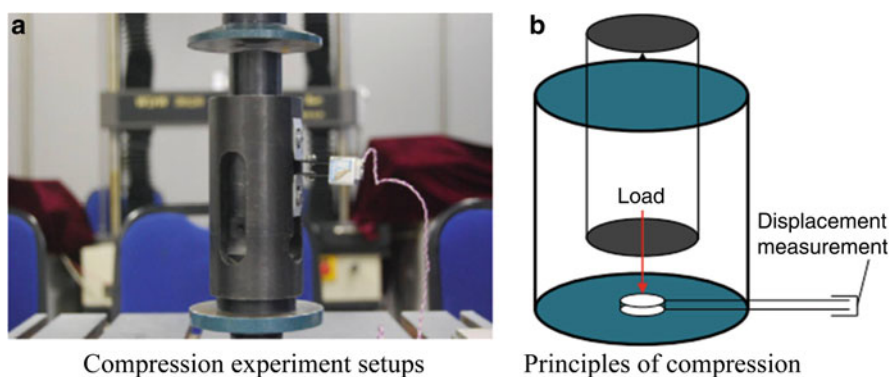


Fig. 23 Quasi-static compression experimental setup on MTS. (a) Compression experiment setups. (b) Principles of compression

Experiment Results

It is commonly accepted that the failure process of composite subject to quasi-static compression loading usually involves a sequential damage accumulation process (Xu et al. 2009). As aforementioned, we set the unloading threshold value to be 80% of the maximum load value due to the major load decrease occurred during the experiment when major crack could be observed clearly. Thus, we define the stress and strain at this moment as major failure on-set (MFO) stress and strain. Figure 24 shows us the “load decrease” phenomenon and the morphology of the cracked samples.

Figure 25 illustrates three different stress–strain relation curves under three loading rates. The stress–strain curves show “nonlinearity” characteristic in mechanical behavior of PVB laminated glass and this “nonlinearity” phenomenon is probably caused by both nonlinear mechanical response of PVB interlayer and progressive

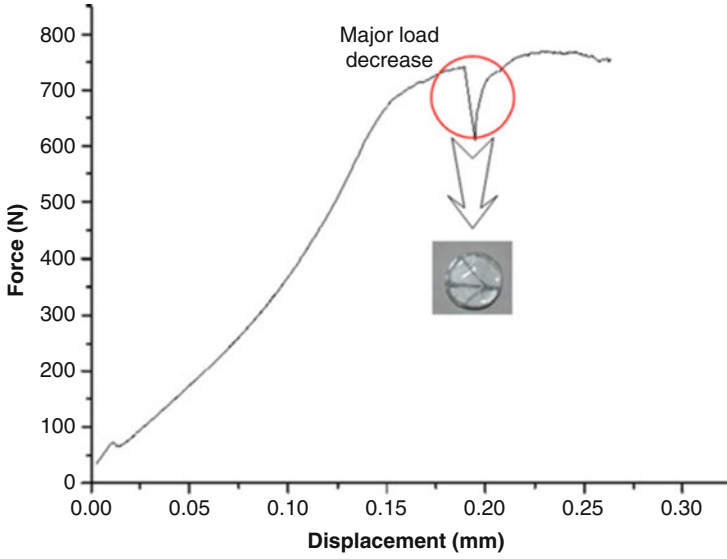


Fig. 24 Major load decrease occurrence with the onset of major crack

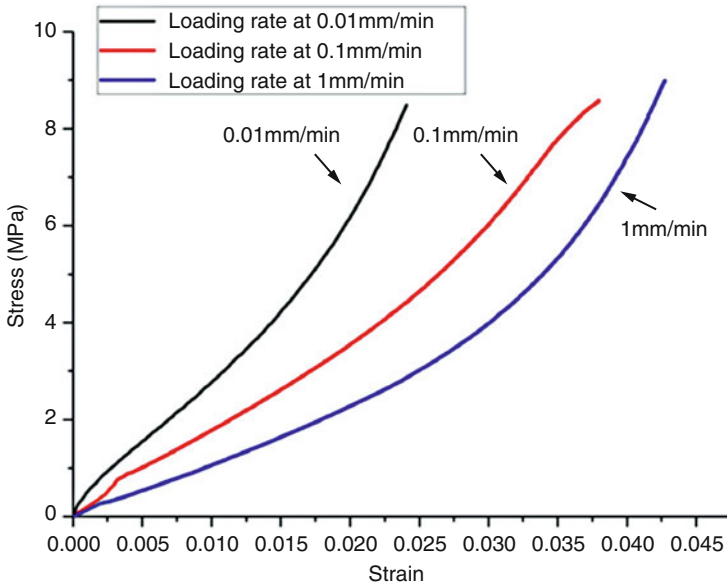


Fig. 25 Stress–strain curves obtained from quasi-static compression experiments under three different loading rates, i.e., 1 mm/min, 0.1 mm/min, and 0.01 mm/min

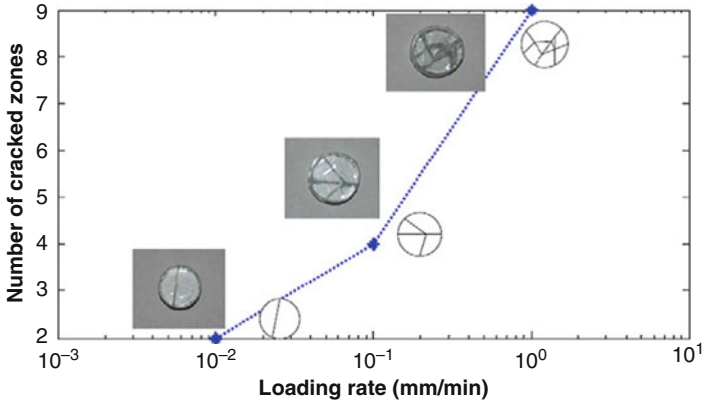


Fig. 26 Number of cracked branching at different loading rates

microcrack growth in sample. As the loading rate increasing, the MFO strain increases while the MFO stress remains nearly the same. The major responsible reasons are: (i) in extremely low strain rate (quasi-static) situation, the outer glass panel plays a critical role in mechanical response and (ii) glass is a rate-independent material whereas PVB is a rate-dependent one.

Number of Crack Branching

The numbers of cracks in testing samples according to different quasi-static loading rates are illustrated in Fig. 26. The number of complete cracks (or major cracked branching) increases with the loading rate. As aforementioned, in quasi-static situation, the stress remains the same whereas the strain rises with the increasing loading rate. Therefore, the number of cracked branching, namely the “damage degree,” depends on the value of strain, not the stress, in quasi-static circumstance that is in accordance with Wang’s (Wang et al. 2006) study. Further, if a nonlinear fitting is tried on the number of crack branching versus loading rate, we will harvest the following explicit formula with fitting coefficient $R = 0.999$.

$$N = [8.975W^{0.339}] \quad (35)$$

where $[X]$ is Round function, e.g. $[5.2] = 5$. N is the number of crack branching and W is loading rate with unit of mm/min. From Eq. 35, we can infer the numbers of crack zones under different loading rates in the quasi-static domain.

Mandelbrot (1977) pointed out that crack branching contained critical information of crack propagation speed and complexity. Bouchaud et al. (2012) further investigated the relationship between crack length and crack branching mode and concluded

$$D_f = \frac{\lg b}{\lg k} \quad (36)$$

where $D_f = \frac{\lg b}{\lg k}$ is degree of branching, b is the number of basic branching pattern, and $K \approx 0.4$ in brittle material.

Assuming that the initial crack branching length is a_0 and the length becomes $a_1 = a_0 + ka_0$ (Bouchaud et al. 2012), thus after the initial crack branches for i times, the length finally comes to

$$a_i = (1 + \dots + K^i) a_0 \quad (37)$$

Since the number of branching can be written as $N \approx \log_b a_i$, we are able to know that

$$i = \log_b^N - 1 \quad (38)$$

According to Freund's (Michel and Freund 1990) theory, the dynamic crack propagation velocity in a finite body obeys the following relation:

$$v = C_r \left(1 - \frac{a_0}{a_i} \right) \quad (39)$$

where C_r is the Rayleigh wave velocity and v is the crack propagation velocity.

By combination of Eqs. 36, 37, 38, and 39, the relation between crack velocity and crack branching number is

$$v = C_r \frac{1 + b^{-1/D_f} - 2N^{-1/D_f}}{2(1 - N^{-1/D_f})} \quad (40)$$

Further, we substitute Eq. 35 into Eq. 40 and get

$$v = C_r \frac{1 + b^{-1/D_f} - 2[8.975W^{0.339}]^{-1/D_f}}{2(1 - [8.975W^{0.339}]^{-1/D_f})} \quad (41)$$

Equation 40 may help us to clarify the crack propagation speeds in different loading rates. We may also conclude that crack propagation speeds are different during the three loading rates in quasi-static experiments based on Eq. 41.

Experiments of Dynamic Out-of-Plane Loading

By starting from a fundamental point, quasi-static mechanical behavior of PVB laminated glass has been widely investigated experimentally above (Xu et al. 2011a). Unlike quasi-static cracking problem, dynamic fracture is more complex and challenging but there are few reported investigations that focused on the dynamic fracture of PVB laminated glass except for some numerical simulations

those based on probabilistic damage mathematical model (Forquin and Hild 2010), continuum damage mechanics (Zhao et al. 2005, 2006a, b; Sun et al. 2005), explicit finite element method (Bois et al. 2003; Timmel et al. 2007; Sun et al. 2009), and extended finite element method (Xu et al. 2010) that often neglected the important PVB layer and most of them did not involve explicit crack growth and pattern analysis. The dynamic experimental investigation is pressing need.

System Setup of Dynamic Out-of-Plane Loading

The impact system setup is sketched in Fig. 27, where a weight block can slide freely along the two standing poles. At the end of the drop-weight, there is a tip which is used to hit the force direction converter on the testing sample and it can be referred in Fig. 28. It is a necessity to convert the vertical impact direction into a horizontal one, which would facilitate the capture of the film.

Upon impact high speed photography is employed for capturing the dynamic crack growth in the plate in situ. The entire system consists of a high-voltage charging controller, multispark box, two field lenses, and an array camera with films (see in Fig. 28). In addition, a multispark high-speed camera is adopted to provide enough spot light source. The multispark box provides 16 independently spot lights triggered by a high-voltage charging controller by a preset adjustable interval time from 1 to 9,999 μs whose charging voltage is able to reach up to 30 kV. Meanwhile, a 4×4 array camera with 16 films is set correspondingly on the other side to receive the light and therefore record the crack image information. In order to make sure that the images recorded on the films are in the time series of the ignition of the spot lights, two field lenses are used to make up the optical path (see Fig. 28b). Moreover, the delay controller is introduced to control the time interval between each spot light. In addition, to minute the exact ignition time for each spot light, the electric-light receiver along with the oscillograph monitor are used together where the electric-light receiver is employed to collect the electric-light signals and send them to the oscillograph monitor. Reference control experiments of impact-induced fracture are carried out using pure glass sheets, with results in good agreement with the literature (Nielsen et al. 2009; Sharon and Fineberg 1999) (see in Fig. 29), thus validating the current test system.

The Preliminary Experiment to Investigate the Radial Crack Propagation

Experiment Condition

In the preliminary impact fracture experiment to investigate the radial crack propagation behavior of PVB laminated glass, the minimum mass of the drop-weight is 0.5 kg and the maximum height of the drop-weight is 1.84 m (corresponds to the maximum impact velocity of 6 m/s); both the mass and height of the drop-weight are also varied in this experimental investigation. Prior to impact, the drop-weight is suspended and fixed through electromagnet. The temperature and relative humidity during the experiments remain in the range of 20–21 and 20–25% respectively.

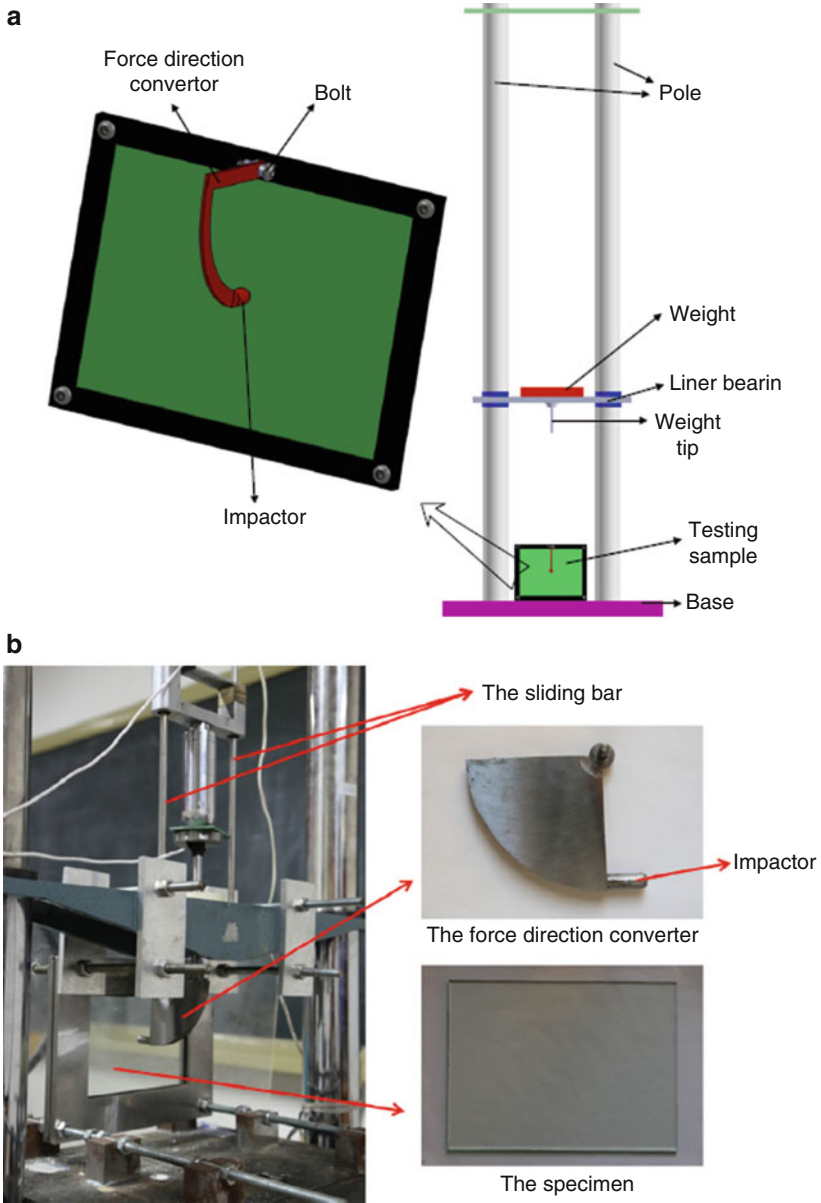


Fig. 27 (a, b) Schematic of the drop-weight tower experiment setup. The testing sample is illustrated along with the force direction convertor

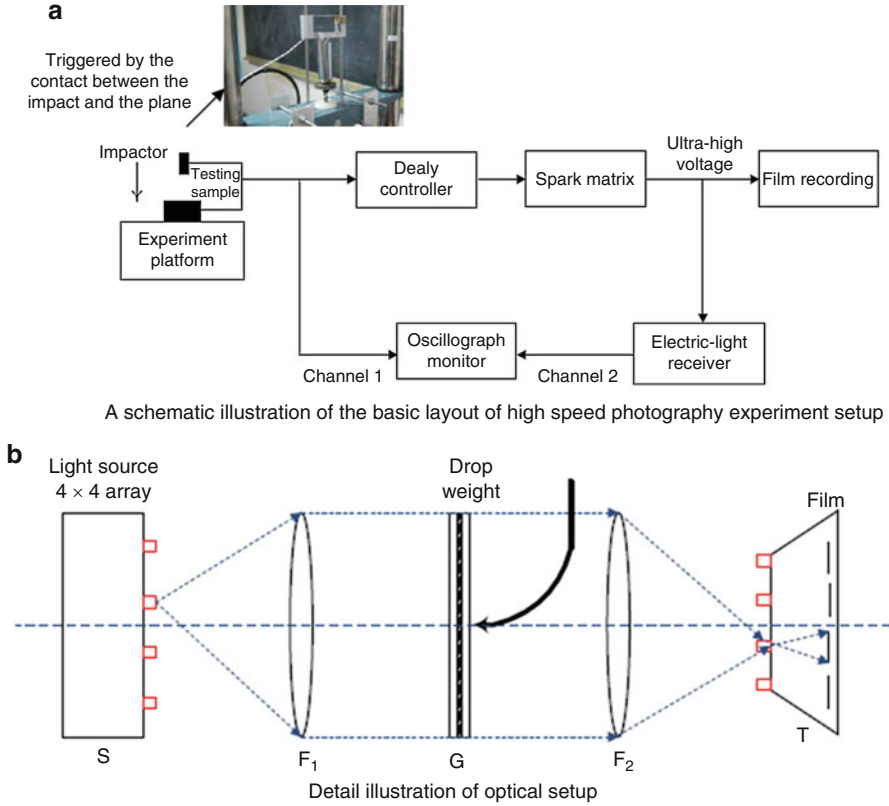


Fig. 28 The schematic illustration of the entire kit of experiment setup. (a) A schematic illustration of the basic layout of high-speed photography experiment setup. (b) Detail illustration of optical setup

Specimen Preparation

The standard PVB layer has a Young’s modulus $E_{PVB} = 100 \text{ MPa}$, Poisson’s ratio $\nu_{PVB} = 0.48$, and mass density $\rho_{PVB} = 870 \text{ kg/m}^3$ (Xu and Li 2009). To make close connection to the automotive industry, the PVB thickness is $t_{PVB} = 0.76 \text{ mm}$ for passenger cars. In addition to the middle PVB layer, for the upper and lower soda-lime glass sheets in laminated glass, the thickness $t_{\text{glass}} = 2 \text{ mm}$, with $\rho_{\text{glass}} = 2,500 \text{ kg/m}^3$, $E_{\text{glass}} = 70 \text{ GPa}$, and $\nu_{\text{glass}} = 0.22$. They are bonded together at 10 bar and 120 °C following standard procedures, and the effective in-plane dimension is $200 \times 150 \text{ mm}$.

Experimental Results and Discussion

Crack Morphology

A series of experiments under different loading conditions, i.e., with different height and mass of the drop-weight are recorded. Figure 30 shows the typical dynamic

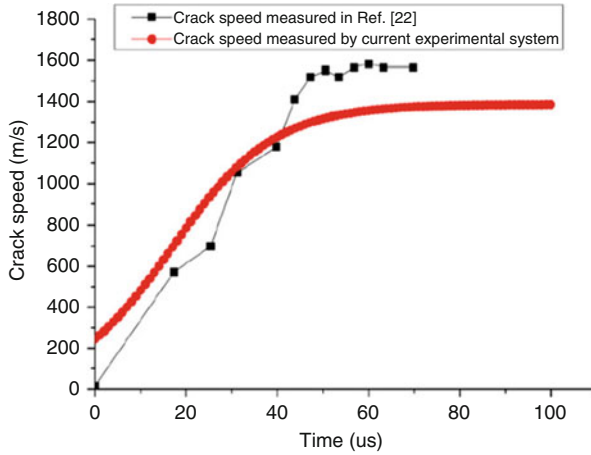


Fig. 29 Comparison of the crack velocity propagation speed of pure glass between experiments results extracted from the current one and the one in Nielsen et al. (2009). Note that the external impact energy is 1 J in the current experiment while it is 0.2 J in the experiment in Nielsen et al. (2009)

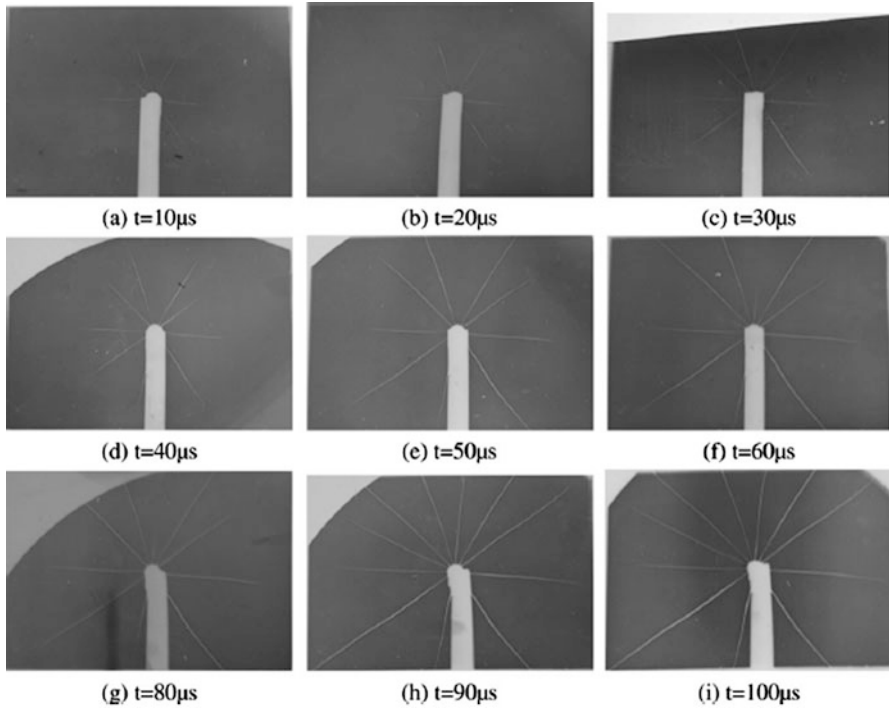


Fig. 30 Snapshots of crack propagation with 1 kg of drop weight and 1 m of drop height

crack propagation snapshots of the PVB laminated glass sheet, in response to a 1 kg weight dropped from 1 m height. Extensive radial fracture is observed, with no appearance of circumferential crack, which is consistent with numerical simulation that when a foreign object impacts a large windshield, the radial crack always appears before the circumferential crack (Xu et al. 2010). It is seen that there are about ten radial cracks on the composite sheet, and the development and evolution of each crack are quite different due to the inevitable stochastic flaw in glass. Here, the average crack length is employed to discuss the behavior of radial crack propagation. By using nonlinear fitting method to the experimental points obtained from the experiments, we finally got the crack tip position in time history. Then the averaged time evolution history of the radial crack tip velocity and acceleration can be obtained from the time history of the radial position of the crack tip as showed in Fig. 31.

The measured crack velocity first increases and then fluctuates slightly, which goes accordance with the experiment results of cracking speeds of pure soda-lime glass at various cracking modes in Nielsen et al. (2009) and Sharon and Fineberg (1999), as well as being consistent with previous numerical simulation results of several brittle materials (Song et al. 2008; Grégoire et al. 2009) and composites (Daphalapurkar et al. 2007; Dwivedi and Espinosa 2003).

As proposed by Ravichandar and Knauss (1984), near the vicinity of crack tip, there is a microcrack zone which impedes the further propagation of crack. Within the impact area, the spacing between the initial cracks is so small and such an intensively damaged zone may reduce the cracking speed substantially. With the

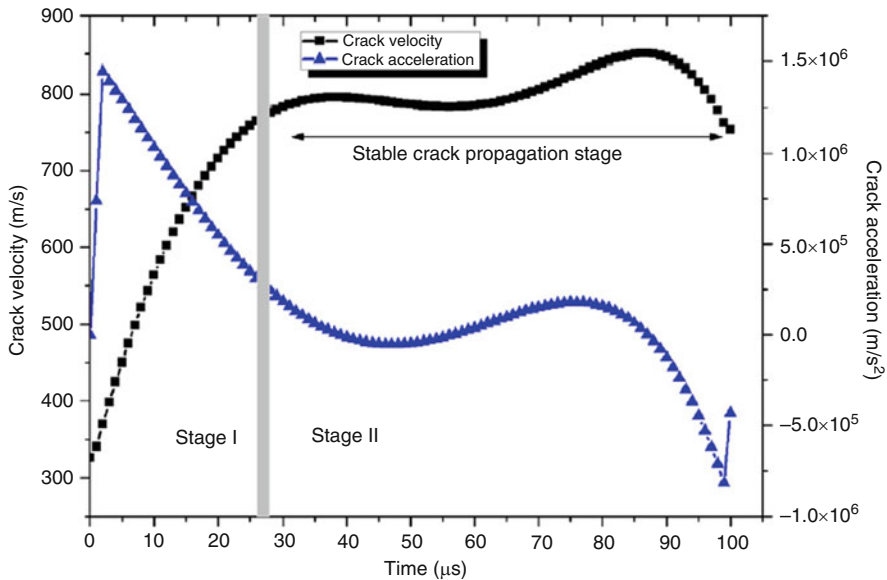


Fig. 31 Averaged crack tip velocity and acceleration in time history for the case in Fig. 28

advancement of cracks, the retardment effect will be greatly lessened until it reaches an almost constant influence, when the crack speed becomes steadier in the later stage. Thus, the crack evolution may be qualitatively divided into two stages: in stage I, the cracking speed keeps rising, and in stage II the crack speed becomes more stable but with some fluctuations. The transition occurs roughly at about time (further discussion will be made later). During stage I, the crack propagation acceleration is decreasing until stage II begins; that is, as the stress in the crack tip accumulates and develops, finally the crack tip will be able to break through the energy barrier surrounding material to reach the equilibrium.

In this example, the averaged steady-state cracking speed during stage II is about $\bar{v}_{PVB} \approx 811 \text{ m/s}$. Comparing with the stable cracking speed of Polymethyl Methacrylate (PMMA) $\bar{v}_{PMMA} \approx 600 \text{ m/s}$ (Sharon and Fineberg 1999) and tempered soda-lime glass $\bar{v}_{GLASS} \approx 1470 \text{ m/s}$ (Nielsen et al. 2009) (One may refer to Fig. 32 for detail comparison. Note that although the impact energy is a little bit different in (Xu and Li 2009) and this experiment, the comparison is still valid), the PVB laminated glass may be regarded a composite material whose overall effective “fracture toughness” is between that of PMMA and pure glass. In Fig. 32, we can also see that the general variation trend of crack speeds in two different materials are almost the same. One may use the Rayleigh wave speed v_R of the glass sheet to normalize the crack speed (since cracking occurs in the glass sheet), i.e., $v_R = 3,098 \text{ m/s}$. Thus, the stable cracking speed of PVB laminated glass is about $0.26 v_R$ in this particular case, which is lower than the typically measured

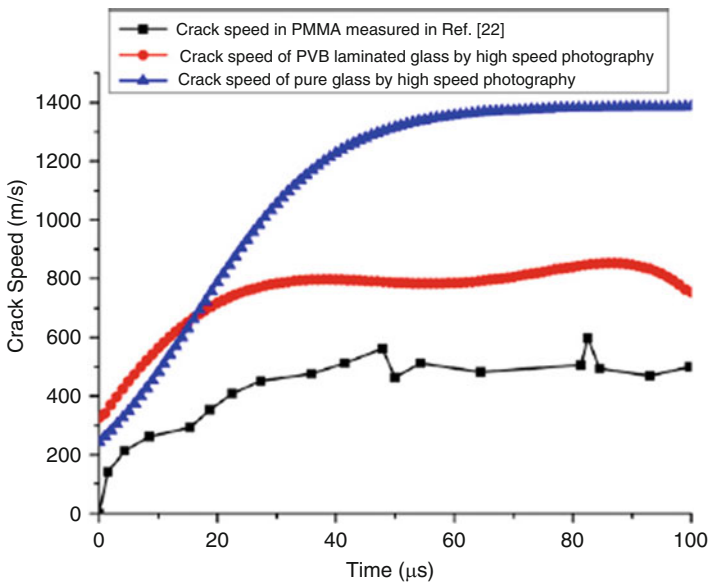


Fig. 32 Comparison of the crack velocity propagation speed among pure glass, PVB laminated glass, and PMMA. Note that data for PMMA is extracted from

maximum speed of about $0.4\text{--}0.5 v_R$ in brittle materials and we can clearly see that the PVB interlayer serves to slow down the crack speed, which would also make the material as well as the brittle-polymer sandwiched structure a candidate for energy absorption/shielding.

Different Driving Mechanisms of In-Plane Cracking on Two Glass Layer

Figure 33 shows a typical crack propagation comparison on supported and loaded glass layers at the loading speed with PVB thickness $h = 0.76$ m. Here, time $t = 0 \mu\text{s}$ is used to denote crack tip initiation for the radial cracks on supported glass layer. It is clear that the radial cracks (Fig. 33a–d) initiated firstly on the “supported glass layer,” then the radial cracks on the “loaded glass layer” generated after about $t = 600 \mu\text{s}$ as shown in Fig. 33f. On the other hand, it is also interesting to observe that the cracks on the two glass layers are completely overlapped even though they propagate at totally different times.

In fact, the stress wave dominated inertial effects during impact loading on the PVB laminated glass play an important role in crack propagation (Singh and Parameswaran 2003; Clements et al. 1996). Due to the high modulus of glass material and the fixed mode of the specimen, the plate is small enough to ignore the bending deformation during the impact. For the convenience of analysis, a schematic diagram in a layer system consisting of two glass plates bonded by a PVB interlayer is shown in Fig. 34. Impact occurs on the surface of the “loaded glass layer” of the PVB laminates, the compressive waves caused by the impact travel

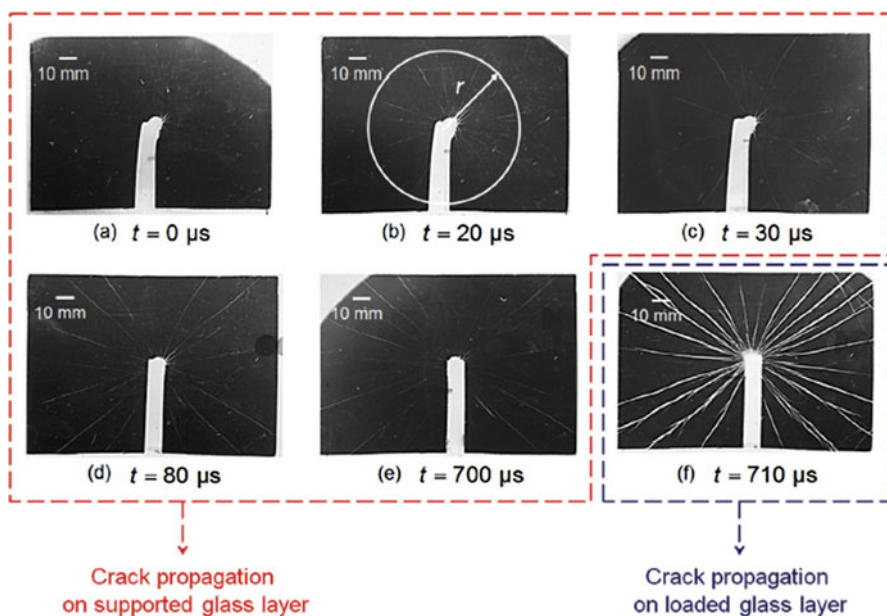


Fig. 33 Selected sequence of images depicting the radial crack growth on both glass layers

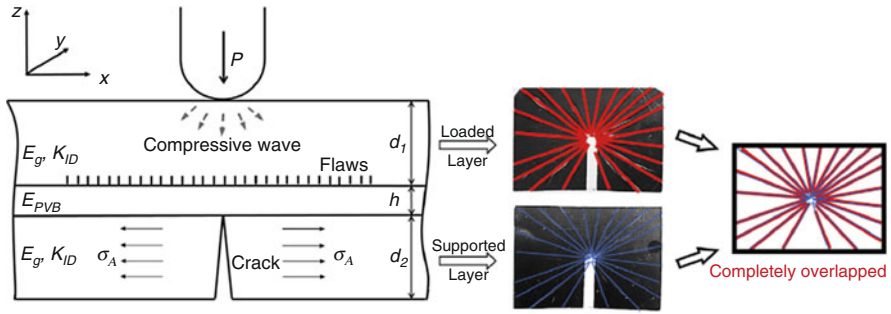


Fig. 34 Schematic of the impact process in a layer system consisting of two glass plates bonded by a PVB interlayer. The cracks propagate into the loaded glass layer by reinitiation from a surface flaw

down the depth direction (Fig. 34). Due to its viscoelasticity, the PVB layer plays an important role in reducing the amplitude and the speed of waves that eventually pass through by internal friction (Thom 2005), compared with the glass layer. Finally, the compressive waves reflect from the boundaries as tensile waves, which continuously arrive at the supported plate thus the tensile dominated cracks initiate while reaching the limit of dynamic fracture toughness KID of the material (Lambros and Rosakis 1997). Considering that glass plate is more sustainable in compressive stress wave but prone to fail in tensile stress wave, the radial cracks initiate first on supported glass layer.

Effect of Drop Weight and Height

Next, we investigate the influence of drop weight and height on the crack propagation. The drop mass ranges from 0.5 to 2 kg, and the height varies from 0.6 to 1.4 m. Table 1 shows the detail combination of the parametric experiments. Figures 7 and 8 show that with the increase of drop weight and height, the basic trend of crack would remain almost the same while the value would increase, i.e., larger mass and higher impact speed will cause faster crack propagation. Zhang et al. (2010) investigated the effect of loading rate on crack velocities in concrete and found with the increase of loading rate, the crack velocity increase proportionally; in quasi-static indentation, it was also found that with larger indentation load, the radial and lateral cracks in ceramic also became longer. All these conclusions are qualitatively consistent with our experimental findings. Particularly, \bar{v}_{PVB} increases 17.70% and 12.21% when the impact speed changes from 3.43 to 4.43 m/s and 5.24 m/s, respectively, and \bar{v}_{PVB} increases 26.72% and 14.92% when the drop weight is increased from 0.5 to 1 kg and 2 kg, respectively. Furthermore, if we combine the effect of impact mass and velocity together into the term of impact energy E , and normalize it by a reference potential energy $E_n = 9.81$ J, then the normalized steady-state cracking speed, i.e., \bar{v}_{PVB}/v_R can be fitted as

$$\bar{v}_{PVB}/v_R = 0.167 \pm 0.137 / (1 \pm 10 \exp(1.50 (0.739 - E/E_0))) \quad (42)$$

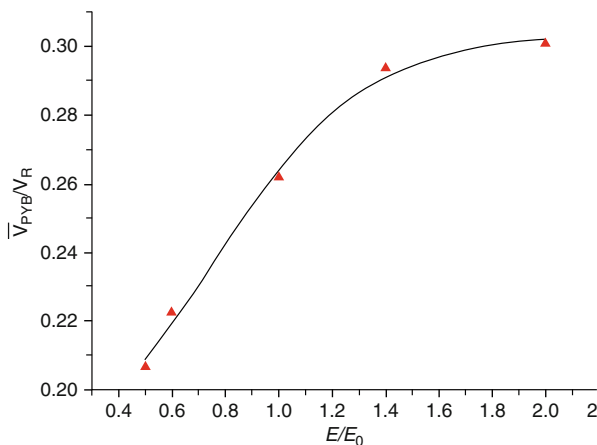


Fig. 35 The relationship between the steady-state crack propagation velocity and the external impact energy

for the range of parameters in this study. From Fig. 35, it can also be found that the stable crack velocity will not likely to increase indefinitely as the external impact energy gets higher, and it has a theoretical limit, i.e., Rayleigh wave speed v_R , but not reaching it.

Regarding the transition time between stages I and II, the general trend is that the smaller the external energy, the faster the transition time (although the difference is relatively small and still on the same order of magnitude). That is to say, if the impact energy is higher, it is easier for the crack speed to achieve steady-state. The detailed physical mechanism of the variation of such a transition time will be explored in future.

The Further Experiment Investigating the Radial and Circular Crack

Experiment Condition

In the further experiment to investigate the radial and circular crack propagation of PVB laminated glass, a weight block with mass of 2 kg is first evaluated to a certain height and then released to create impact loading. The height of the drop-weight adopted in our work varies from 300 to 900 mm. An impactor is fixed at the tip of the force direction converter to control the contact shape and area. Different shapes may be designed for the impactor, and in this study, the impactor shape is confined to be cylindrical and the top is hemispherical. The diameter of the cylindrical is 10 mm and the total length of the impactor is 25 mm. The mass of the impactor (including the force direction converter) is 0.545 kg. In addition, a force sensor with sufficient accuracy and frequency is attached right below the drop-weight to record force–time history and the signals recorded by the force sensor are sent to the oscillograph monitor to record force–time history.

Specimen Preparation

The plate is consisted of a PVB interlayer sandwiched by two brittle glass sheets (For PVB interlayer: Young's modulus $E_p = 0.1$ GPa, short time shear modulus $G_0 = 0.33$ GPa, long time shear modulus $G_1 = 0.69$ MPa, bulk modulus $K = 20$ GPa, Poison's ratio $\nu_p = 0.49$, density $\rho_p = 1,100$ kg/m³, and decay factor $\beta = 12.6$ s⁻¹; For glass: Young's modulus, Poison's ratio and thickness of the glass. $E_g = 70$ GPa, Poison's ratio $\nu_g = 0.22$ and thickness of the glass is 2 mm. The thickness of the PVB interlayer varies from 0.38 to 3.04 mm during our parametric experimental study (where thickness of 0.76 mm is for common passenger cars). The specimen is prepared using the same manufacturing process (with compression at 10 bar and 120 °C) as that used in automotive windshield. Glass is a typical linear elastic brittle material where its facture strain is about 0.1% while the PVB is a rubber-like material which may sustain a much larger deformation during tension. The unique design enables the possible protection against impact with small amount of energy and dissipates larger impact energy through the large deformation of interlayer.

The specimen is clamped within two metal cover sheets with thin layers of rubber pad inside to avoid possible scratches and stress concentration on the sample surface and distribute the boundary force more uniformly. The tightening torque for the screw bolt is set to be 4 Nm which properly mimics the clamping boundary condition. Six screws penetrating the metal cover sheets are fixed to make sure the metal cover sheets are placed in right horizontal position and reduce the rebound under impact. A force direction conversion part (indicated in Fig. 27) is employed to transfer the vertical impact force into a horizontal one, providing an out-of-plane dynamic loading to the sample.

Experimental Results and Discussion

Crack Morphology

Two kinds of macroscopic crack patterns, radial crack and circular crack, are observed and all the pieces of glass plates after impact are stay connected to the PVB layer. Hoop stress is larger than radial stress such that radial crack appears before circular crack, therefore, it gives us an opportunity to capture and study the evolution history of radial and circular cracks separately. Figure 36 is typical crack propagation process which contains a series of selected sequence of images depicting the radial and circular crack growth in PVB laminated glass specimen respectively. Time $t = 0$ μ s is used to denote crack tip initiation for both radial and circular cracks. The crack propagation velocity v is determined as the crack length increase Δl over the time increase Δt . For radial crack, at time t_i , multiple major crack tips conform approximately a circle with radius r_i on average, shown in Fig. 36a. Therefore, radial crack propagation velocity $v_r = \Delta l / \Delta t = (r_{i+1} - r_i) / (t_{i+1} - t_i)$. For circular crack, the total length of n main cracks at t_i is added and divided by the crack number as \bar{l}_i and such that circular crack propagation velocity $v_c = \Delta l / \Delta t = (\bar{l}_{i+1} - \bar{l}_i) / (t_{i+1} - t_i)$. In addition,

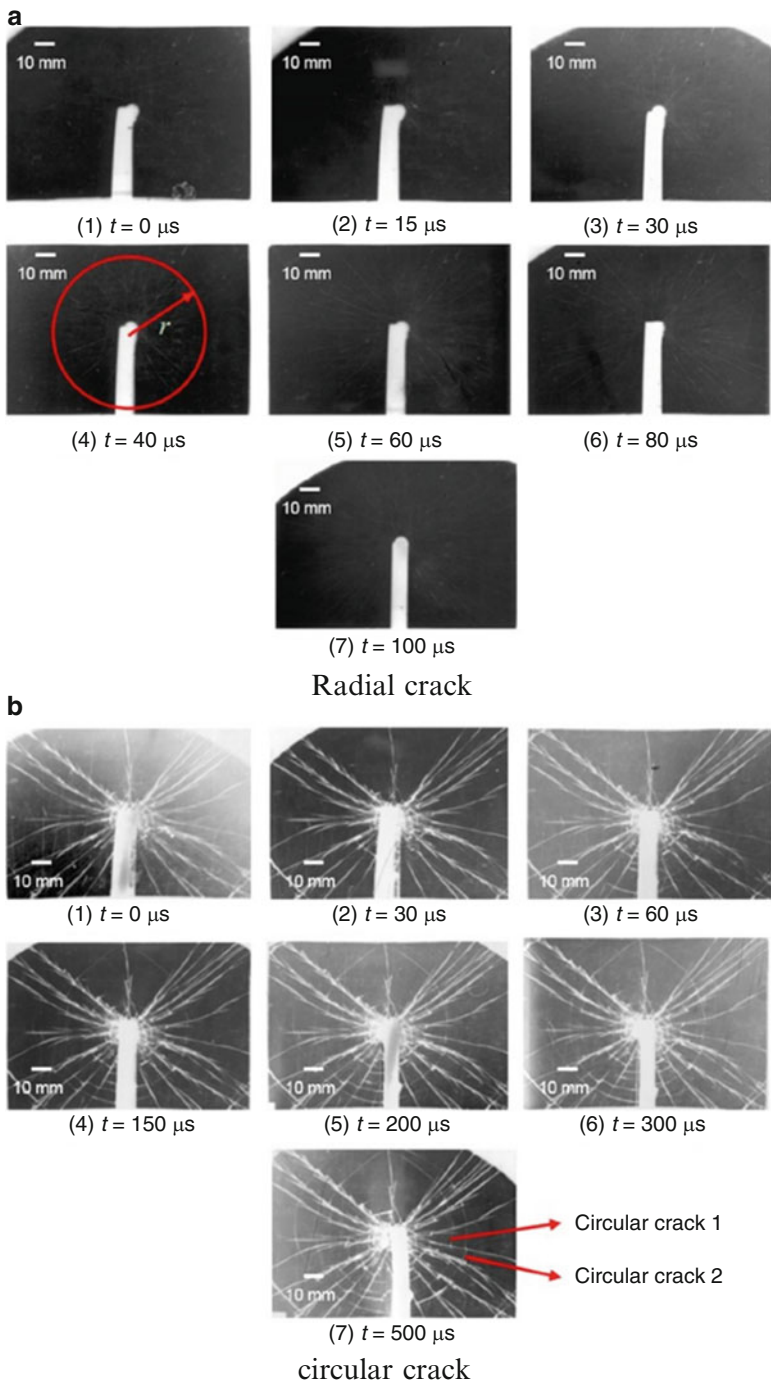


Fig. 36 Selected sequence of images depicting the crack growth at the impact velocity of 3.13 m/s, with PVB thickness of 0.76 mm. (a) Radial crack. (b) Circular crack

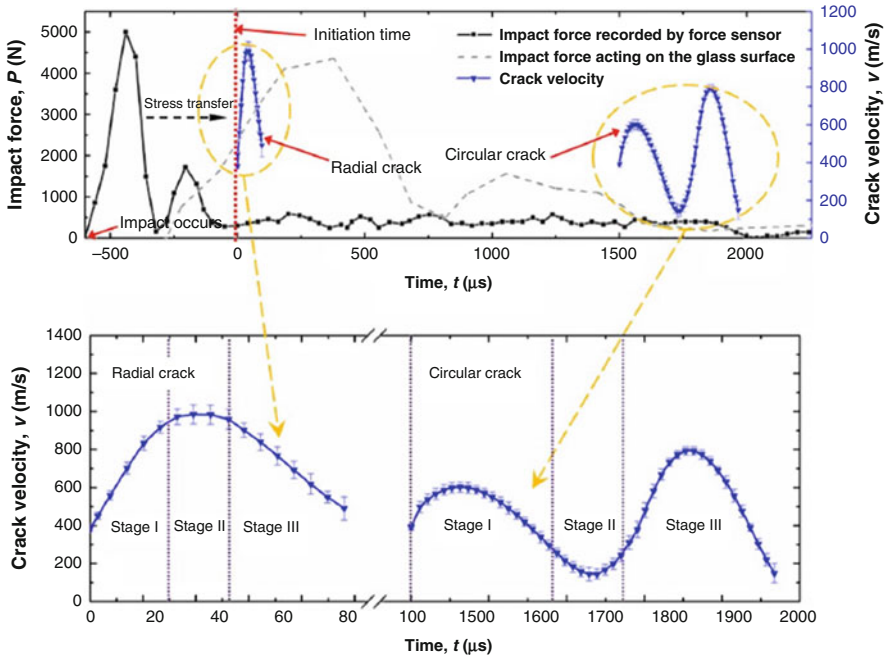


Fig. 37 Load–time curve and crack growth velocity history of the radial crack and the circular crack, in response to 500 mm drop height (impact velocity:3.13 m/s) and 0.76 mm PVB thickness

the B-spline interpolation method is used to smooth the experimental points in this study.

Crack Propagation

To enhance the credibility of the experimental results, the dynamic strain indicator is introduced. Locations near the force sensor and the top of the impactor are fitted with strain gages respectively so as to monitor in situ strains (and stresses) transmit during the impact process. Thus, we could estimate the delay of the impact force from the force sensor to the glass surface and depict the accurate force–time curve acting directly on the glass surface (see dashed line in Fig. 37). Figure 37 shows a typical crack propagation velocity time history curves for radial and circular crack propagation velocities respectively in response to 2 kg weight from 500 mm drop height (impact velocity: 3.13 m/s) to a specimen with 0.76 mm PVB interlayer. Time $t = 0 \mu\text{s}$ is employed to denote crack tip initiation for radial crack. Therefore, the negative time correspond to preinitiation conditions (see Fig. 37). Note that the crack only exists in the glass sheets such that the stress component normal to the crack front σ_{11} is the major driving force for crack initiation and propagation. The dynamic crack propagation velocity v may be expressed as

$$v^2 = \frac{\sigma_{11} + \sum_{11}}{\rho_g (1 + \epsilon_{11})} \tag{43}$$

where the contact stress at the crack tip which can be calculated based on existing dynamic contact model by attaching a strain gage near the impactor. \sum_{11} is the strain component. By supposing the proportionality coefficient D between σ_{11} within the framework of linear elasticity (Zhang et al. 2010), the above equation may be rewritten as

$$v^2 = \frac{1 + D}{\rho_g (1/\sigma_{11} + 1/E_g)} \quad (44)$$

In this impact loading condition, the maximum crack propagation velocity is 990 m/s and 790 m/s for radial and circular cracks, respectively. As Xu et al. (2010) indicated that σ_{11} is larger in hoop stress than circumferential stress so radial crack propagates faster than circular crack. The radial crack propagation velocity is lower than that in pure soda-lime glass (Nielsen et al. 2009) where no obvious circular crack occurs. The PVB interlayer acts as an energy-absorbing material thus to reduce the crack propagation velocity. Meanwhile, PVB interlayer provides a connection media for cracked glass panel such that circular crack could continue to dissipate impact energy after the full growth of radial cracking.

Impact occurs on the outer surface of the glass plate at around $-250 \mu\text{s}$ (see Fig. 37). A compressive loading wave front travels down at around 5,300 m/s (Zhao et al. 2006b) in the glass layer. The PVB layer presents the dual advantages of delaying the passage of the elastic wave into the backing glass layer and reducing the amplitude and the wavelength of wave that eventually passes through by internal friction caused by its viscoelasticity. Both of these effects are beneficial in utilizing the energy absorbing capabilities of the PVB interlayer to the fullest. The compressive waves arrive at the free surface of the backing layer and reflect as tensile waves. Thus, the delay of the crack initiation probably results from the transmitting and reflection of the stress wave and the increasing process of the stress intensity factor caused by increasing elastic energy stored near the crack tip. Similar phenomena have also been observed in Lambros and Rosakis (1997). More and more tensile waves continually arrive at the backing plate and thus the radial cracks initiate while reaching the limit of dynamic crack initiation toughness K_{IC}^d . For the rate-sensitive material, the PVB interlayer is rate dependent while the fracture occurs in the glass plate which is rate independent. Considering the fact that cracks grow in the glass layer, therefore, the impact velocity has little effect on the K_{IC}^d value, which means the effect of rate sensitivity can be ignored. The propagation of the radial crack can be divided into three stages (see Fig. 37). After initiation, the crack tip accelerates from about 400 m/s to around 990 m/s within $30 \mu\text{s}$. We can see that the impact force remain increasing after the initiation of the radial crack, which means the elastic energy stored near the crack tip keeps increasing thus causing the acceleration sequences. Then the crack velocity remains constant at about 990 m/s during the next $25 \mu\text{s}$. However, it decelerates to less than 600 m/s in under $100 \mu\text{s}$ after initiation. The entry of crack tip into the region of compressive stress as it approaches the boundaries and the arrival of less stress wave at the crack tip are two responsible reasons for the deceleration phenomenon (Lambros and Rosakis 1997).

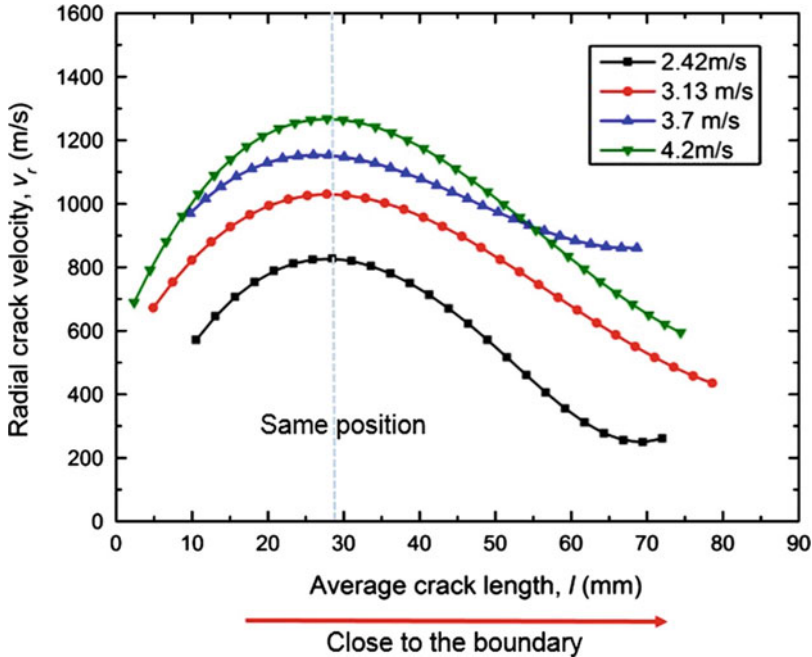


Fig. 38 Velocity–length curve of the radial crack at various impact velocities from 2.42 to 4.2 m/s with the PVB thickness of 0.76 mm

Through further investigation, the reason for the deceleration and crack arrestment phenomenon is relevant to the boundaries. The history of the radial crack velocity at various impact velocities from 2.42 to 4.2 m/s with the PVB thickness of 0.76 mm is plotted as a function of crack length in Fig. 38 and it is clearly shown that the crack velocity start decreasing at the same position.

For the circular crack, the crack initiation occurs 1,750 μ s after the impact on the glass surface (see Fig. 37). Combined with the images depicting the radial and circular crack growth, it is believed that circular cracks initiate long after the radial crack growth as we stated before. An observation of the specimen, at the end of the test, showed that circular cracks merely existed on the loaded layer, according to which we can conclude that the circular cracks nuclear under the influence of surface acoustic waves (namely “Rayleigh waves”). The propagation of this kind of elastic waves is limited to the loaded surface of the solids. The Rayleigh waves spread at a much lower speed compared with the longitudinal compression waves. Besides, the lateral dimensions of the sample are much larger than those in longitudinal direction. Therefore, the circular crack appears generally later than the radial crack which is confirmed by the previous finding (Xu et al. 2010). The circular crack propagates under the tensile stress caused by the reflected tensile Rayleigh waves, which results in that the circular cracks always emerge first close to the boundaries. The propagation of the circular crack can be divided into three stages as well (see

Fig. 37). Both the first and the third stage contain the acceleration, the maintenance, and the deceleration process similar as those in radial cracks. There is more space for circular crack to grow than radial crack such that the total crack lengths are longer in circular cracks. Besides, the circular crack has the lower peak value. However, between the two stages, there exists a special stage which could be called “slowing-down stage.” In this stage, the crack propagation velocity decelerates to almost 0. This is a direct consequence of the fracture at the glass interface caused by the radial crack. We can clearly see that the circular crack propagation always delays at the relatively large fracture where the stress waves have to propagate through the interlayer, without which the circular crack cannot transfer effectively across the interface (Park and Chen 2011).

Effect of Impact Velocity

To investigate the effect of impact velocity v_d , the both radial and circular crack propagation characteristics, respectively, series of parametric experiments are conducted. Figure 7 shows the crack propagation velocity time history of both radial and circular cracks under various impact velocities from 2.42 to 4.2 m/s (drop height from 300 to 900 mm). Two kilogram drop weight and 0.76 mm PVB thickness are employed in this study. Three repeated experiments under each condition are carried out to ensure the results accuracy (see error bars in Fig. 39). Time $t = 0 \mu\text{s}$ is

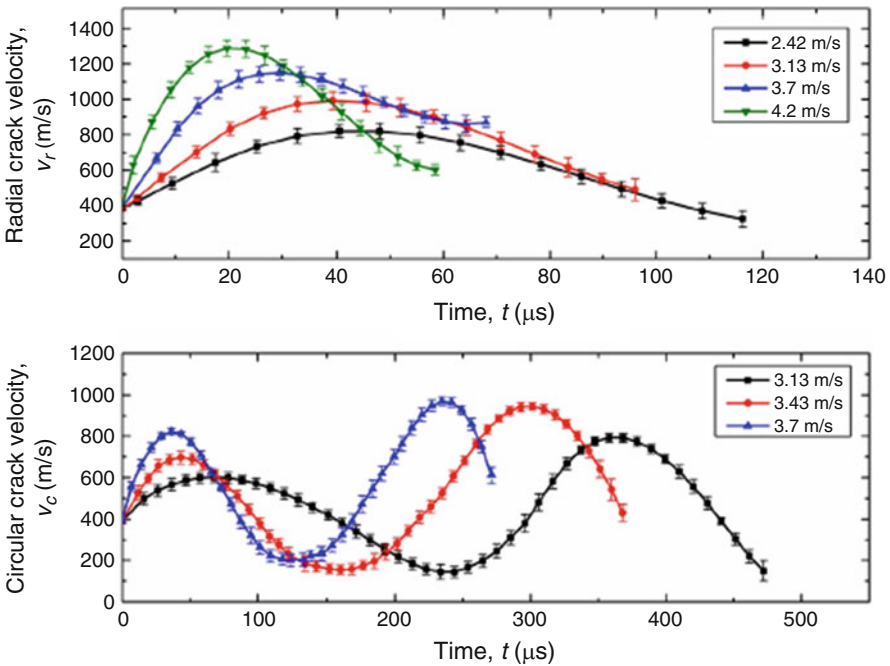


Fig. 39 Crack velocity history of the radial and circular crack at various impact velocities from 2.42 to 4.2 m/s with the PVB thickness of 0.76 mm

employed to denote crack tip initiation. For the radial crack, in the case with higher impact velocity, a shorter acceleration time is needed to reach the top propagation velocity where the top value is also higher. This is consistent with the fact that at higher impact velocity with larger wave amplitude, the load applied to the specimen increases at a faster rate and thus crack acceleration can be larger during the dynamic loading process. Besides, additional experiments also indicate that the higher the impact velocity, the shorter delay time is needed before initiation (see Fig. 40). However, the case with higher impact velocity would also decelerate faster and earlier due to the limited available cracking space in the panel plane. Similar to the radial crack, higher impact velocity results in the shorter initiation time, the shorter acceleration time, and the larger peak value for the circular crack. To establish the model combining the crack velocity with the impact conditions, we chose the variable “the maximum velocity” to quantify our study, which apparently depends on both the impact conditions and the material property. We can suppose that the maximum velocity v_{\max} at the crack tip during the whole propagation process has an implicit relationship with the impact force, the duration of the loading time and the dynamic initiation fracture toughness of the material in the present investigation.

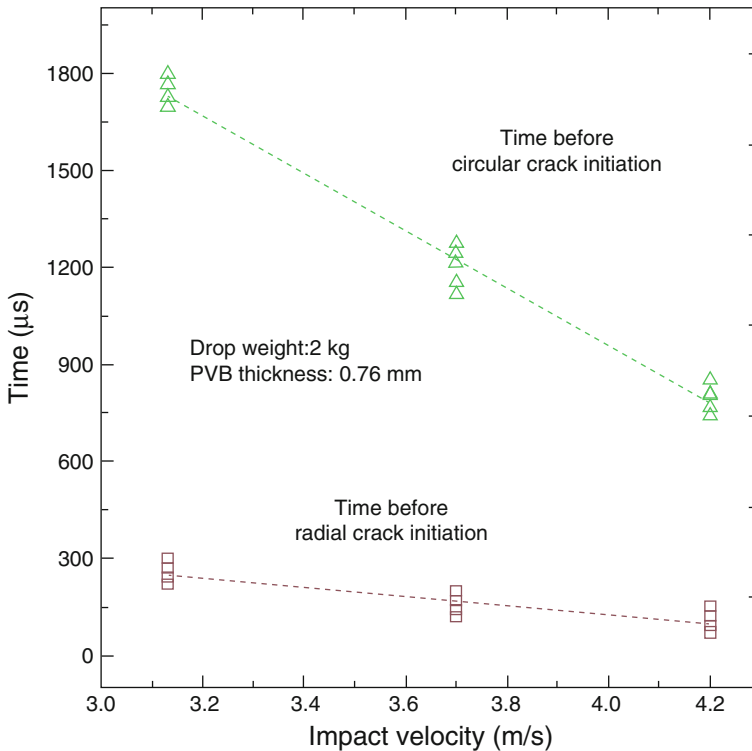


Fig. 40 Initiation time of the radial and circular crack at various impact velocities from 3.13 to 4.2 m/s with the PVB thickness of 0.76 mm

According to the principle of dimensionless unifying, the equation of v_{\max} could have the following form as:

$$v_{\max} = \frac{AP_d^{2/3}}{(k_{IC}^d)^{2/3} t_d} \quad (45)$$

where p_d is the average contact force over the contact period and t_d is the duration of the loading, respectively, A is a macroscopic crack pattern dependent coefficient. According to the momentum theorem, we have:

$$p_d t_d = m_d v_d \quad (46)$$

This means that the maximum velocity at the crack tip can be expressed in terms of the impact velocity v_d and the drop weight m_d

$$v_{\max} = \frac{Am_d^{2/3} v_d^{2/3}}{(k_{IC}^d)^{2/3} t_d^{5/3}} \quad (47)$$

Thus, the maximum velocity contains two model parameters: the dynamic initiation fracture toughness k_{IC}^d , which can be geometry dependent, and the macroscopic crack pattern dependent parameter $A > 0$. Both parameters are subject to the experimental data via best-fitting.

Effect of PVB Thickness

Series of parametric experiments are conducted to study the effect of PVB thickness t_p on the crack propagation characteristics as well. Three different PVB thicknesses, i.e., 0.38 mm, 0.76 mm, and 1.52 mm are chosen for the parametric experiments. Two kilogram drop weight is employed. The crack propagation velocity time history of both radial and circular cracks under three different PVB thicknesses can be seen in Fig. 41. The impact velocity used to investigate the radial crack is 2.42 m/s (i.e., drop height is set to be 300 mm) while the impact velocity is 3.13 m/s (i.e., drop height is set to be 500 mm) for circular crack. Martinez et al. (1998) reported that the capability for transmitting the impact energy depends on the thickness and type of the adhesive used. Therefore, the variation in PVB thickness will have a great effect on delaying the passage of the elastic wave into the backing glass layer and reducing the amplitude and the wavelength of wave through internal friction (Xu et al. 2011b), which results in reducing the maximum radial crack speed as show in Fig. 41. Considering the wave speed in the through thickness direction of around 5,320 m/s in the glass layer and 300 m/s in the PVB interlayer (which means the PVB interlayer plays an important role in slowing down the stress wave), the transit time is around 3 μ s, 6 μ s, and 11 μ s, respectively with the PVB thickness ranged from 0.76, 1.52, to 3.04 mm. Compared with the delay

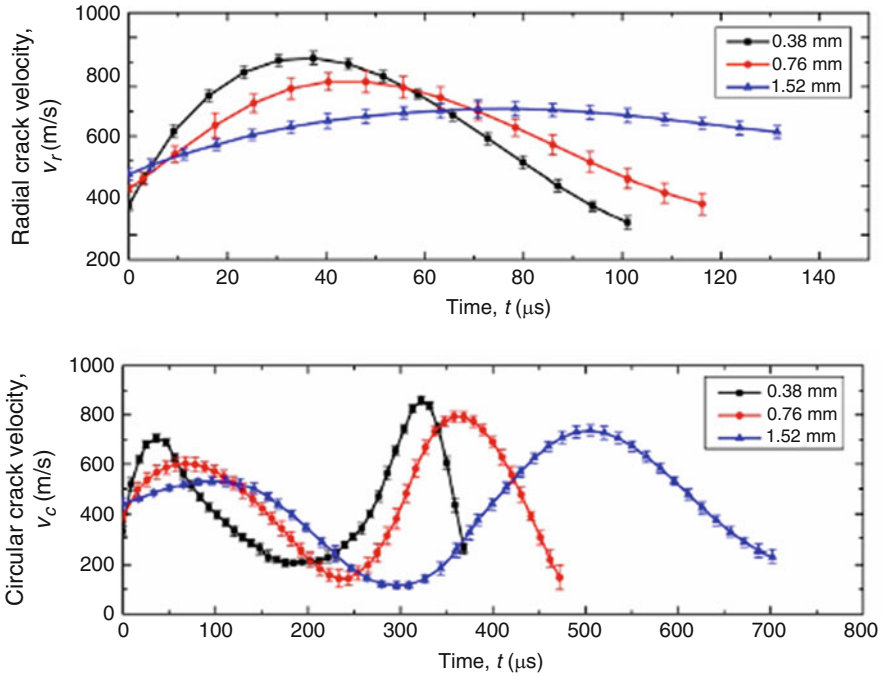


Fig. 41 Crack velocity history of the radial and circular crack with various PVB thicknesses from 0.38 to 1.52 mm at the impact velocity of 2.42 m/s and 3.13 m/s, respectively

time in the order of 102 μ s for radial crack initiation in Fig. 10, we can conclude that the delay of the radial crack initiation mainly results from the increasing process of the stress-intensity factor. The thicker PVB thickness would reduce the amplitude of elastic wave crossing the PVB interlayer by internal friction thus to weaken the increasing process of the stress-intensity factor. Therefore, in the case with thicker PVB layer, a longer delay time is needed to initiate. For the circular crack caused by Rayleigh wave, similar phenomenon is observed. Considering the Rayleigh surface wave speed in the in-plane direction of around 3,370 m/s in the glass layer (Sharon et al. 2002), the transit time is around 37 μ s, which is much longer than the transmit of the elastic wave in the medium. Similarly, the delay of the circular crack initiation mainly also results from the increasing process of the stress-intensity factor. In addition, from the force–time curve, we can see that the compressive waves stay around the loaded side for a relatively long time during the impact, which also have an effect on the increasing process of the stress-intensity factor.

It is clearly seen that the glass sheet with thicker PVB layer starts at a higher initiation velocity v_0 for both crack patterns. Since the exact expression of v_0 is not clarified yet, we may implicitly represent the expression to the initiation velocity v_0 as $v_0(t_p)$. Due to the fact that $k_{IC}^d \propto \sigma_c \propto v_0$ (see Eq. 44), where σ_c is the fracture

stress, we have the following expression for the initiation velocity (Berezovski and Maugin 2007).

$$v_0^2(t_p) = c_R^2 (1 + D) \left(1 - \left(1 + \frac{(k_{IC}^d)^2}{M} \right)^{-1} \right) \quad (48)$$

where M is a constant coefficient and C_R is the Rayleigh wave velocity. In this case, we arrive at an expression for the dynamic initiation fracture toughness in Eq. 47 in the form of

$$k_{IC}^d = \sqrt{\frac{M v_0^2(t_p)}{C_R^2 (1 + D) - v_0^2(t_p)}} \quad (49)$$

Therefore, we may come to the expression for the maximum velocity at the crack tip as follows:

$$v_{\max} = \left[\frac{B m_d^2 v_d^2 t_d^{-5} (c_R^2 (1 + D) - v_0^2(t_p))}{v_0^2(t_p)} \right] \quad (50)$$

where $B > 0$ is another constant coefficient depending on the macroscopic crack pattern. According to Eq. 50, the thicker the PVB layer, the lower of the maximum velocity for both radial and circular crack, which is consistent with the experiment result (see Fig. 41). This is expected since the glass sheet with thicker PVB layer would absorb more impact energy via molecule vibration and friction by the interlayer as polymer-like material, verified by Xu et al. (2011c). As the Rayleigh wave velocity $C_R \approx 3370$ m/s (Hauch and Marder 2010), which is sufficiently high for the initiation velocity v_0 . The expression may be simplified as

$$v_{\max} = \left[\frac{B m_d^2 v_d^2 t_d^{-5} (c_R^2 (1 + D))}{v_0^2(t_p)} \right]^{1/3} = c \left[\frac{m_d^2 v_d^2 t_d^{-5}}{v_0^2(t_p)} \right]^{1/3} \quad (51)$$

Then the agreement between theory and experiment may be achieved by an appropriate choice of the values of the constant coefficient c (c_1 and c_2 for radial and circular crack, respectively). The fitted models for the maximum velocity of radial crack v_{\max}^r and circular crack v_{\max}^c are in the following form of:

$$\begin{cases} v_{\max}^r \approx c_1 \left[\frac{m_d^2 v_d^2 t_d^{-5}}{v_0^2(t_p)} \right]^{1/3} = 0.0657 \left[\frac{m_d^2 v_d^2 t_d^{-5}}{v_0^2(t_p)} \right]^{1/3} \\ v_{\max}^c \approx c_2 \left[\frac{m_d^2 v_d^2 t_d^{-5}}{v_0^2(t_p)} \right]^{1/3} = 0.0545 \left[\frac{m_d^2 v_d^2 t_d^{-5}}{v_0^2(t_p)} \right]^{1/3} \end{cases} \quad (52)$$

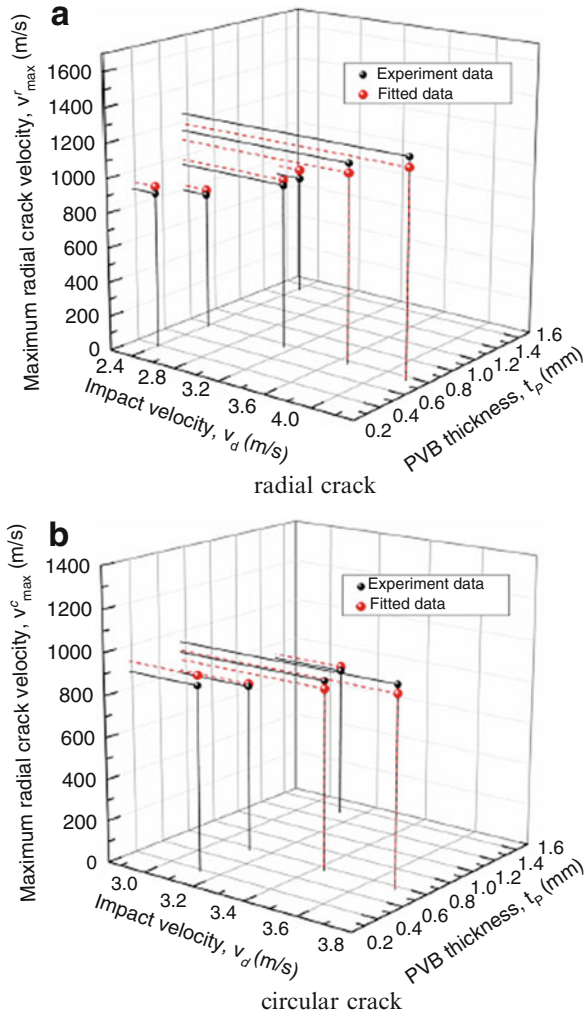


Fig. 42 Experiment data and fitted data of the maximum crack velocity from the semophysical with various impact velocities from 2.42 to 4.2 m/s with PVB thickness from 0.38 to 1.52 mm. (a) Radial crack. (b) Circular crack

The fitted data and experimental data are shown together in Fig. 42 (the experimental data in the figure is the average maximum crack velocity). The good agreement between the two proves the model to be dependable.

Statistical Analysis on Macroscopic Cracking Morphology

To confirm the crack propagation mechanism analysis before and understand the crack propagation mechanism from another perspective, the relationship between

Table 4 Parametric statistical experiments: linear regression estimates of the Weibull parameters under different experiment conditions

Crack type	PVB thickness (mm)	Impact velocity (m/s)	Mean crack number	Weibull char. number (N_0)	Weibull modulus (m)
Radial crack	0.38	2.42	87	96	4.4
	0.76	2.42	81	89	4.0
	1.52	2.42	70	79	3.1
	3.04	2.42	66	75	2.7
	0.76	3.13	90	99	4.4
	0.76	3.7	99	107	5.1
	0.76	4.2	118	128	6.0
Circular crack	0.38	2.42	2.5	2.7	2.8
	0.76	2.42	2	2.2	2.5
	1.52	2.42	1.7	2.0	2.2
	3.04	2.42	1.5	1.6	1.7
	0.76	3.13	2.7	3.1	2.7
	0.76	3.7	3.9	4.2	5.1
	0.76	4.2	7.0	7.4	9.3

the number of cracks and the impact velocity or the PVB thickness is established. Since all cracks locate in the surface of glass sheet and glass is a material with internal stochastic flaws which may serve as crack nuclei, a statistical model may be needed to describe the two macroscopic crack patterns. Table 4 shows the detail combination of the statistical experiments. One hundred repeated experimental data for each specified condition have been used for each experiment condition to obtain a statistical result. The two-parameter Weibull model is employed to study the distribution of both radial and circular crack number under each experiment condition. Weibull statistics are commonly used in the engineering community, especially in the fracture statistics. In the specific case of Weibull statistics analysis in this study, the cumulative probability function is written as follows:

$$p_f = 1 - \exp\left(-\left(\frac{N}{N_0}\right)^m\right) \quad (53)$$

where P_f is the cumulative probability, N is the number of radial or circular crack (N_r or N_c), m and N_0 is the distribution shape parameter and characteristic crack number, respectively. To improve the statistical accuracy, the number of circular crack is counted with decimal during the statistical process (that is the number of circular crack with incomplete circle is determined by the length ratio of incomplete circle and the complete one). For the ease of accessing information, the double logarithm of the Weibull Eq. 55 is used in number analysis, yielding the Weibull parameters in a simple graphical representation of the data. Figure 43 shows the

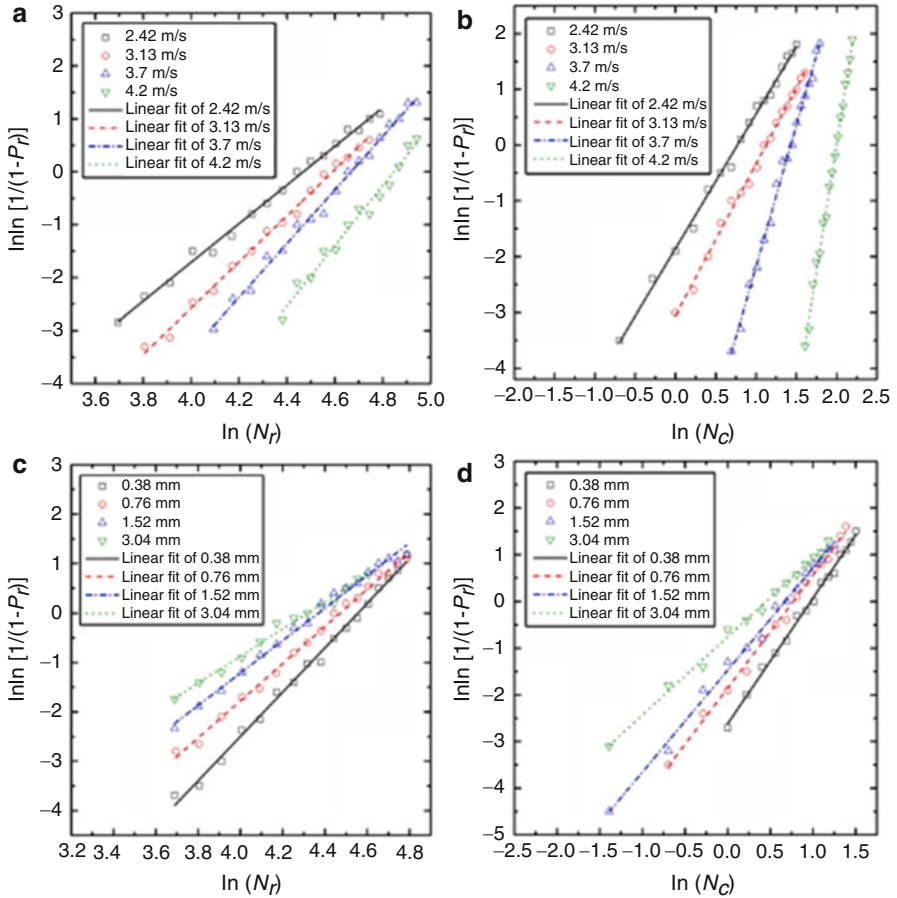


Fig. 43 Fitted Weibull model plots for macroscopic cracking morphology: (a) radial crack at various impact velocities from 2.42 to 4.2 m/s; (b) circular crack at various impact velocities 2.42–4.2 m/s; (c) radial crack with different PVB thicknesses from 0.38 to 3.04 mm; and (d) circular crack with different PVB thicknesses from 0.38 to 3.04 mm

graphs with $X = 1 - p_f$ plotted on the horizontal axis and $Y = \ln[\ln(1/(1 - p_f))]$ on the vertical axis. One can perform a simple linear regression analysis to get the Weibull parameters under different experiment conditions. The slope of the line is the Weibull modulus, m . The characteristic crack number, N_0 is the value of crack number for which $\ln[\ln(1/(1 - p_f))]$ is zero. In this study, the linear regression analysis is used for the Weibull parameter estimates under each experiment condition (see Table 4).

$$\ln(1 - p_f) = -\left(\frac{N}{N_0}\right)^m \tag{54}$$

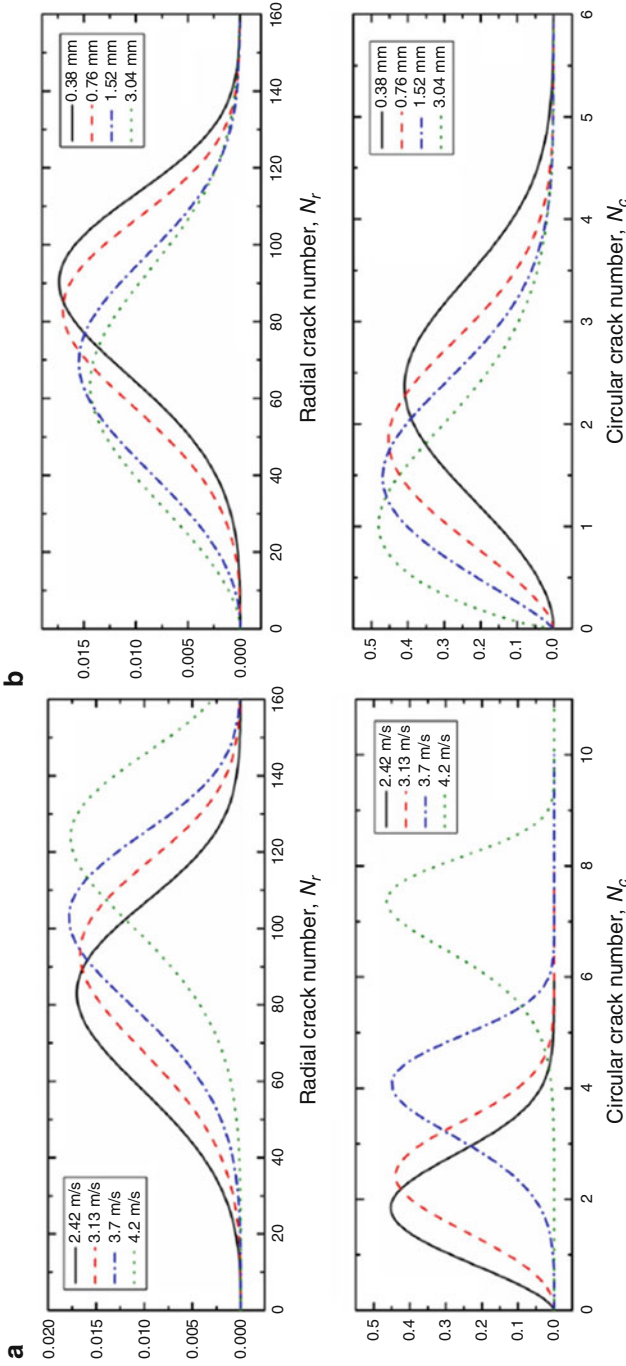
$$\ln \left[\ln \left(\frac{1}{1 - p_f} \right) \right] = m \ln N - m \ln N_0 \quad (55)$$

Firstly, the effect of different impact velocities on the distribution of both radial and circular crack number is investigated. Figure 44a shows the probability density curves of radial and circular crack number under different impact velocities. The characteristic crack number indicates the distribution location along the X axis and is expected to vary according to the impact velocity or the PVB thickness. The distributions for both radial and circular crack number are proved to move to the right (higher value) as impact velocity increases. This is reasonable because higher impact velocity produces more elastic energy, and thus it is necessary to create larger new surface to release the excess energy. In addition, the probability density curves considering different PVB thicknesses to describe the radial and circular crack number are shown in Fig. 44b. Contrary to the effect of impact velocity, the distributions for both radial and circular crack number move to the left (lower value) as PVB thickness increases since thicker PVB layer absorbs more elastic energy as stated before. This statistical method could also be applied to the engineering applications and play an important role in the laminated glass design.

The Relation Between Radial Crack Velocity and Crack Numbers on the Backing Glass Layer

Here, the effect of the radial crack number on their own crack propagation velocity as well as the crack generation on the other glass layer is thoroughly investigated. Firstly, the radial crack velocity on the backing glass layer with different crack numbers is calculated. A theoretical model from the perspective of energy conversion is established to depict the cracking process of backing glass layer and elucidate the fundamental reason that causes the variation of the crack velocity.

The specimen in response to impact speed $V = 3.7$ m/s is chosen for 100 repeated experiments by considering the intrinsic stochastic flaws in glass laminated samples. A series of images depicting the crack growth on both glass plates are recorded at the set time intervals. Typical crack propagation process is shown in Fig. 45. The glass plate directly contacting with the impactor is defined as “impacted glass layer,” while the glass plate on the other side is referred as “backing glass layer.” $t = 0$ μ s represents the time when impact is triggered. Figure 45a–d demonstrate the whole cracking process on the backing glass layer where cracks always initiate first, and only radial crack pattern is observed while both radial and circular cracks appear on the impacted glass layer (under the effect of Rayleigh waves, whose propagation is limited to the loaded surface of the solids) long after the cracking of the backing one (Fig. 45e–f). However, the radial crack number on the backing glass plate varies from 18 to 112 within 100 repeated experiments, which is quite a large deviation. Further, it is also qualitatively observed that the number of radial



For the radial and circular crack number with four different PVB thicknesses, i.e. 0.38 mm, 0.76 mm, 1.52 mm and 3.04 mm.

For the radial and circular crack number at four different impact velocities, i.e. 2.42 m/s, 3.13 m/s, 3.7 m/s, and 4.2 m/s

Fig. 44 Probability density curves of the radial and circular crack number under various boundary conditions. (a) For the radial and circular crack number at four different impact velocities, i.e., 2.42 m/s, 3.13 m/s, 3.7 m/s, and 4.2 m/s. (b) For the radial and circular crack number with four different PVB thicknesses, i.e., 0.38 mm, 0.76 mm, 1.52 mm, and 3.04 mm

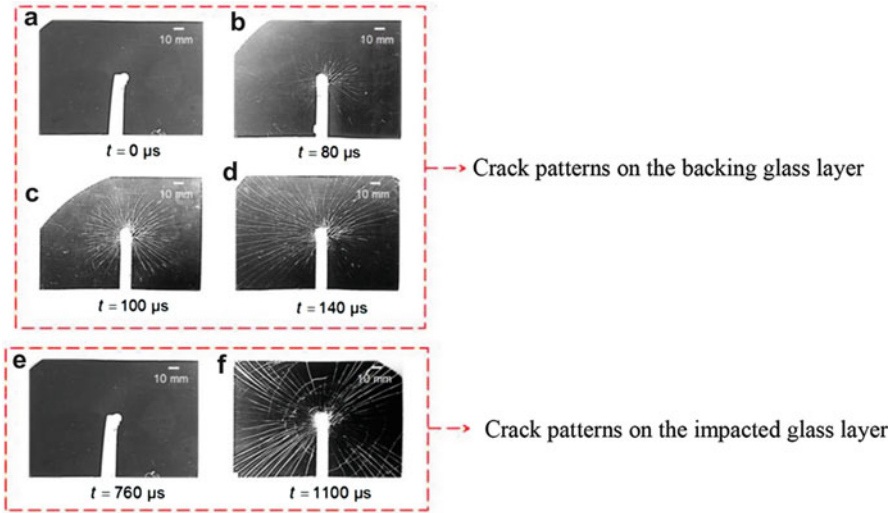


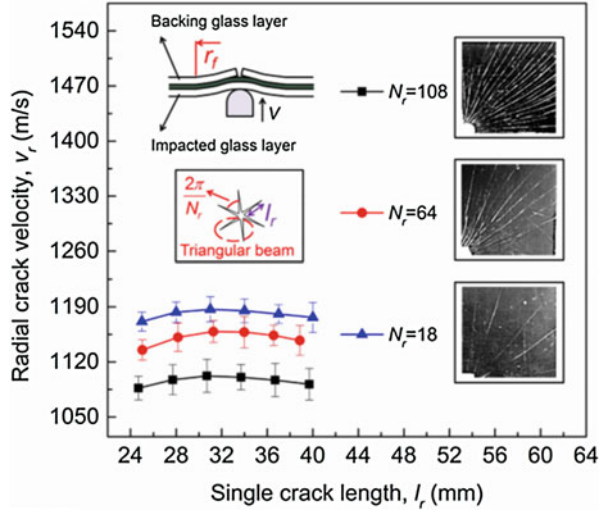
Fig. 45 Images depicting the crack growth on the two glass layers ($V = 3.7$ m/s)

cracks have great effect on the later crack generation on the impacted glass layer as well as the propagation characteristics of radial crack on the backing glass layer itself. In the following section, the effect of the different radial crack numbers of backing glass layer would be thoroughly discussed.

The crack velocity history on the backing glass layer with different radial crack number is calculated separately. All the radial cracks present approximately the same length during the propagation, conforming a circle overall (Fig. 45). Therefore, the crack velocity v_r is calculated as the radius increase Δr over the time increase Δt , i.e., $\Delta r/\Delta t$. Figure 46 shows the crack velocity variation with the crack length l_r under three selected radial crack numbers ($N_r = 18, 64,$ and 108). Crack propagations far from the boundaries and the center of the glass plate are focused to eliminate the possible boundary effect (i.e., the crack length between 20 and 40 mm). Results indicate that the number of radial cracks on the backing glass layer regularly negatively influences cracking propagation velocity themselves, i.e., the crack velocity v_r always decreases with the radial crack number increasing.

During the propagation of the cracks on the backing glass layer, the impacted glass layer remains intact as stated before (Fig. 46). However, due to the adhesion by PVB interlayer, the two glass layers have the same radius of deformed region r_f (Fig. 46). The bending energy of the impacted glass plate caused by the impact can be estimated by assumption that the backing glass layer is still intact, in which case the two glass plates can be regarded as “single one with $2h_g$ thickness.” Therefore, the bending energy of impacted glass layer is equal to half the energy in the single glass plate with $2h_g$ thickness. After the contact between the impactor and the glass plate, kinetic and bending energies balance $\rho_g \Omega V^2 \sim E_g (2h_g \kappa)^2 \Omega$ (Vandenberghe et al. 2013), where $\Omega = 2h_g \pi r_f^2$ is the volume of deformed region and V is the impact

Fig. 46 Radial crack velocity versus crack length with different radial crack number ($V = 3.7$ m/s)



speed whose variation can be ignored during such a short period. Thus, the curvature $\kappa \sim w_0/r_f^2 \sim V/2h_g c_g$, yielding $r_f \sim (2h_g c_g t)^{1/2}$, where $w_0 = V_t$ is the indentation and $c_g = \sqrt{E_g/\rho_g}$ is the sound speed in the glass material. The bending energy of impacted glass plate U_{b1} can be expressed as

$$U_{b1} = k_1 \pi E_g \left(\frac{V}{c_g} \right) r_f^2 h_g \quad (56)$$

where k_1 is a coefficient.

Considering the low Young's modulus of PVB material (i.e., $E_p = 0.1$ GPa), the bending energy of PVB U_{b3} is about 0.001 of that in the glass plate, which has been proved by the finite element simulation with a three-dimensional laminated plate FE model using 1×1 mm surface size element. The boundary condition of the glass plate model is set to be clamped, which is consistent with the real experimental condition. Thus, U_{b3} is small enough to be ignored compared to that of U_{b1} .

While for the backing glass layer in the cracking process with crack length l_r , the number of radial cracks N_r on the backing layer has been set and would not change as the cracks extend. Thus, the bending energy of the backing glass plate U_{b2} can be estimated by the energy of N_r triangular beams of length l_r with the neglect of transverse bending (Vandenbergh et al. 2013):

$$\begin{aligned} U_{b2} &= E_g h_g^3 w_0^2 N_r \tan\left(\frac{\pi}{N_r}\right) / 3l_r^2 \\ &\approx k_2 \pi E_g h_g r_f^2 \left(\frac{V}{c_g}\right)^2 \left(\frac{1}{2} + \frac{\pi^2}{6N_r^2}\right) \end{aligned} \quad (57)$$

where k_2 is another coefficient. Therefore, the bending energy of the whole laminated plates U_b can be estimated as

$$U_b = (U_{b1} + U_{b2}) \approx \pi E_g h_g r_f^2 \left(\frac{V}{c_g}\right)^2 \left[k_1 + k_2 \left(\frac{1}{2} + \frac{\pi^2}{6N_r^2} \right) \right] \quad (58)$$

From the perspective of energy transferring, during the cracking process, the elastic energy of the laminated plates (i.e., the bending energy) caused by the impact is converted into the glass surface energy through the fracture behavior. Here we define the increasing rate of the plates elastic energy as “ η_b ,” and the energy release rate is “ η_c .” Therefore, we have

$$\eta_b = \frac{dU_b}{dt} \approx \pi E_g h_g^2 \frac{V^2}{c_g} \left[k'_1 + k'_2 \left(\frac{1}{2} + \frac{\pi^2}{6N_r^2} \right) \right] \quad (59)$$

$$\eta_c = \frac{d(2N_r \Gamma h_g l_r)}{dt} = 2N_r \Gamma h_g v_r \quad (60)$$

where k'_1 and k'_2 are both coefficients. Γ is the glass material fracture (surface) energy. There should be a balance between the “input energy” (i.e., the increasing elastic energy) and the “output energy” (i.e., the fracture energy). Suppose that the proportion between the two parameters η_b/η_c , which is named as “energy conversion factor,” indicating the stability of the system or the ability for further crack growth, then the larger proportion would obviously lead to a greater degree of instability for the whole system. Thus for the system of PVB laminated plates in our study, the “energy conversion factor” η_b/η_c can be expressed as

$$\frac{\eta_b}{\eta_c} = \frac{\pi E_g h_g}{2\Gamma c_g} \cdot V^2 \cdot \frac{1}{v_r} \left\{ \frac{1}{N_r} \left[k'_1 + k'_2 \left(\frac{1}{2} + \frac{\pi^2}{6N_r^2} \right) \right] \right\} \quad (61)$$

One may clearly see that the expression of the “energy conversion factor” contains three parts: (1) $\pi E_g h_g / 2\Gamma c_g$ is determined by the material properties (here refers to the glass material), (2) V^2 refers to the impact condition, and (3) rest of the expression $f(v_r, N_r)$ refers to the impact responses of the laminates. Thus the stability of the system is determined together by the material properties, impact condition, and the impact response of the material according to Eq. 61. Previous studies (Xu et al. 2011d; Chen et al. 2013) investigating the effect of the impact speed on the crack velocity have demonstrated that, with higher impact speed V that increasing the system instability η_b/η_c , both of the crack velocity v_r and the radial crack number N_r tend to increase as well thus to lower the value of η_b/η_c such to maintain the balanced energy input and output in a large extent. Therefore, as the crack number N_r on the backing glass layer decreases under the same impact condition, the value of η_b/η_c would correspondingly increase due to the result of Eq. 61, which means the instability of the system is rising. Thus the system tends

to response in a way to reduce this instability (i.e., the value of η_b/η_c) through propagating the cracks at higher speed, improving the value of v_r in Eq. 61, which is consistent with our experiment results (Fig. 46).

The Relation Between the Crack Length and the Capability of Energy Absorption on the Impacted Glass Layer

Further, the total crack length for both two crack patterns on the impacted glass layer is calculated to study the capability of energy absorption of laminated glass influenced by the radial crack number on the backing glass layer.

Figure 47 shows the crack morphology on the impacted glass layer where both of the two crack patterns (i.e., radial crack and circular crack) appear long after the cracking of the backing glass plate. For the radial cracks on the impacted layer that are indirectly caused by the stress concentration brought by each generated crack on the backing glass layer as initial flaw (Lee et al. 2007) (which is also proved by our experiment investigation), they are completely overlapped with the radial cracks on the backing layer (Fig. 47). Thus, it is obvious that the morphology as well as the number of radial cracks in the impacted glass layer depends completely on that of the backing layer. The radial cracks on the impacted glass initiates from those on the backing glass plate through the depth direction (Lee et al. 2007).

Specimen in response to lower impact speed $V = 2.42$ m/s is chosen for 100 repeated experiments. As expected, the radial crack number on the backing glass plate varies from 15 to 65 in these experiments. The total crack length for both two patterns on the impacted glass layer is calculated with different crack numbers on the

Fig. 47 Selected sequence of images depicting the crack patterns on both glass layers at the loading speed $V_d = 3.7$ m/s, with PVB thickness $h = 0.76$ mm

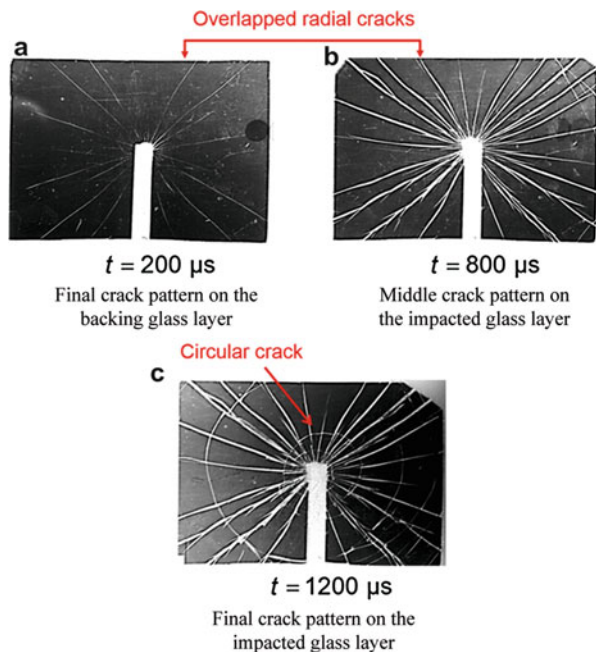
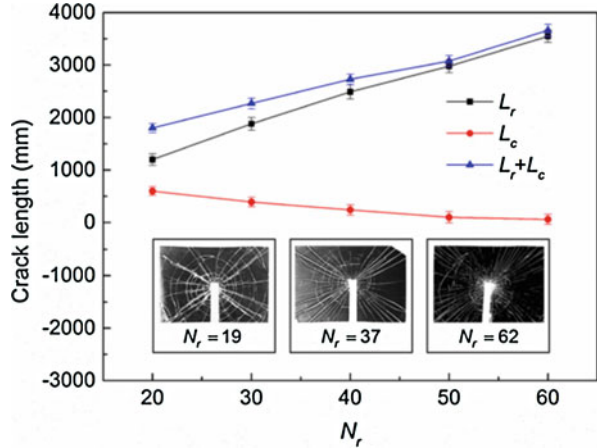


Fig. 48 Crack length on the inner glass layer versus radial crack number ($V = 2.42$ m/s)



backing glass layer (Fig. 48). L_r and L_c are defined as the total radial and circular crack length on the impacted glass layer, respectively. Thus, $(L_r + L_c)$ refers to the total crack length on the impacted glass plate. The radial crack length linearly increases with crack number on the backing glass layer due to the uniform growth of all cracks while the variation trend of circular crack length is decreasing (Fig. 48). Circular cracks always initiate long after the completed propagation of radial cracks on the impacted glass (Xu et al. 2010) (Fig. 47). Thus, before the initiation of circular cracks, the laminated plates can be regarded as N_r triangular beams with thickness $2h_g$ and length l_0 (which can be estimated as the side length of specimen) since the radial cracks on the two glass plates are completely overlapped (Fig. 47). The bending energy U_{b0} of the plates is

$$U_{b0} = 4\pi E_g h_g^3 w_0^2 (1 + \pi^2/3N_r^2) / (3l_0^2) \quad (62)$$

In the case with less radial crack number N_r , the bending energy would increase, resulting in stronger stress concentration in the radial direction, which makes the later generation of circular cracks easier. In addition, taking the law of energy conservation into account, the total fracture energy on the whole laminated plates can be estimated as $2\Gamma L h_g$, where L is the total crack length on the two glass plates. Thus, the fracture energy released on two glass plates can be estimated by calculating the crack length on each plate. For the plates with shorter radial crack length, leaving more residual energy, it will benefit the generation of circular cracks as well.

Although the increase of circular crack number contributes positively to the energy absorption of the laminated plates, the effect is much less compared with the radial cracks (Fig. 48). As the radial crack morphology on the impacted glass is directly determined by that on the backing glass layer, it is safe to conclude that the backing glass plate is the key that decides the fracture morphology and further

the capability of energy absorbing of the whole laminated structure such that the material selection and structure design for the backing layer is extremely important for impact protection.

Summary

In this chapter, we have investigated the propagation characteristics for both radial and circular cracks in PVB laminated glasses by theoretical constitutive equations analysis, numerical simulation, experiments, and tests of impact.

Firstly, a damage-modified nonlinear viscoelastic constitutive relations model of PVB laminated windshield is suggested based on updated Lagrangian method to study the crack evolutions on windshield plate subject to pedestrian head.

Then, the constitutive relations are implemented into FEA software to simulate the pedestrian head impact with vehicle windshield after verified by a classical example and classical Hertzian pressure calculation model. Based on the FE model above, the internal stress of PVB laminated glass is analyzed and found that shear stress, compressive stress, and tensile stress are main causes of plastic deformation, radial cracks, and circumferential cracks, respectively for the laminated glass subject to impactor. By a parametric study based on the FE model above, the increase of Poisson's ratio will lead the increase of crack angle. Impact velocity has a great influence of crack formations as well as the injuries of pedestrian head. Comparing with other numerical techniques, XFEM is arguably more efficient since it does not require remeshing during crack propagation. So the extended finite element method (XFEM) is adopted to numerically investigate the multiple crack propagation of PVB laminated glass subject to quasi-static indentation or low-speed impact. It is found that the critical accident information, such as the impact speed or damage stress, can be deduced from the crack pattern characteristics. A qualitative bridge can be established between numerical simulation result and real-world accident via the crack growth mechanism.

Lastly, for the experimental analysis of PVB laminated glass, crack branching based on crack fractal theory is investigated and an explicit expression describing the crack velocity and number of crack branching is proposed under quasi-static Split Hopkinson Pressure Bar (SHPB) compression experiments. For dynamic out-of-plane loading, we carried out preliminarily impact fracture experiments firstly to investigate the radial crack propagation behavior of PVB laminated glass subjected to light-weight impact by using high-speed photography. The time histories of the averaged radial crack tip position, propagation velocity, and acceleration are obtained. It is found that the steady-state cracking speed of PVB laminated glass is lower than that of pure glass, and it also increases with higher impactor speed and mass. It is also found that the supported glass layer would always initiate before the loaded layer and the final morphologies of radial cracks on both sides are completely overlapped even if they propagate at different time. Here, two different mechanisms of crack propagation on different glass layers are explained: the cracks on supported glass layer are dominantly motivated by the in-plane crack-tip stress

such that a relatively longer time is needed for its propagation while for the cracks appearing much later on loaded layer, which are caused by the stress concentrated in depth direction for each generated crack on supported layer, the propagation time needed is largely reduced. Then, the further experiment investigating the radial and circular crack propagation of PVB laminated glass thoroughly is carried out. The parametric study on two dominant factors, i.e., impact velocity and PVB thickness is discussed aiming at the crack propagation of PVB laminated glass. Results show that cracking speeds of both radial crack and circular crack increase with higher impact velocity. Radial crack appears earlier than circular one and it propagates at higher speeds. Based on the mechanism of crack initiation and propagation, a semiphysical model describing the relationship between the maximum cracking velocity and influential factors is established. In addition, the Weibull statistical model is suggested considering various factors to describe the macroscopic crack pattern thus providing a theoretical evidence for engineering practice. Finally, the relation between radial crack velocity and crack numbers on the backing glass layer and the relation between the crack length and the capability of energy absorption on the impacted glass layer are investigated. Results show that the radial crack velocity on the backing glass layer decreases with the crack number under the same impact conditions during large quantities of repeated experiments. Thus, the “energy conversion factor” is suggested to elucidate the physical relation between the cracking number and the crack propagation speed. However, the fracture energy of the whole laminated plates is mainly determined by the radial crack length, which means that the radial crack number generated first on the backing glass plate is the key that decides the capability of energy absorbing of the whole laminated structure.

References

- A. Berezovski, G.A. Maugin, On the propagation velocity of a straight brittle crack. *Int. J. Fract.* **143**, 135–142 (2007)
- P.A.D. Bois, S. Kolling, W. Fassnacht, Modelling of safety glass for crash simulation. *Comput. Mater. Sci.* **28**, 675–683 (2003)
- E. Bouchaud, J.P. Bouchaud, J.P.S.G. Lapasset, The statistics of crack branching during fast crack propagation. *Fractals-Compl. Geom. Patterns Scaling Nat. Soc.* **1**, 1051–1058 (2012)
- J. Chen, J. Xu, X. Yao, B. Liu, X. Xu, Y. Zhang, Y. Li, Experimental investigation on the radial and circular crack propagation of PVB laminated glass subject to dynamic out-of-plane loading. *Eng. Fract. Mech.* **112–113**, 26–40 (2013)
- B.E. Clements, J.N. Johnson, R.S. Hixson, Stress waves in composite materials. *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdisc. Top.* **54**, 6876–6888 (1996)
- N.P. Daphalapurkar, H. Lu, D. Coker, R. Komanduri, Simulation of dynamic crack growth using the generalized interpolation material point (GIMP) method. *Int. J. Fract.* **143**, 79–102 (2007)
- S.K. Dwivedi, H.D. Espinosa, Modeling dynamic crack propagation in fiber reinforced composites including frictional effects ☆. *Mech. Mater.* **35**, 481–509 (2003)
- F. Zhigang and Z. Jianping, On the viscoelastic finite element method, *Shanghai J. Mech.* **16**(1), 20–26 (1995)
- U. Fischer, K. Washizu, *Variational Methods in Elasticity & Plasticity*, 3rd edn. (Pergamon Press, Oxford/New York, 1982.) XV, 630 S., £ 40.00. US \$ 100.00. ISBN 0 08 026723 8, Zamm

- Journal of Applied Mathematics & Mechanics Zeitschrift Für Angewandte Mathematik Und Mechanik, **64** (1984) 70–71
- P. Forquin, F. Hild, A probabilistic damage model of the dynamic fragmentation process in brittle materials. *Adv. Appl. Mech.* **44**, 1–72 (2010)
- D. Grégoire, H. Maigre, A. Combescure, New experimental and numerical techniques to study the arrest and the restart of a crack under impact in transparent materials. *Int. J. Solids Struct.* **46**, 3480–3491 (2009)
- J.A. Hauch, M.P. Marder, Energy balance in dynamic fracture, investigated by a potential drop technique. *Int. J. Fract.* **90**, 133–151 (2010)
- J. Lambros, A.J. Rosakis, Dynamic crack initiation and growth in thick unidirectional graphite/epoxy plates. *Compos. Sci. Technol.* **57**, 55–65 (1997)
- B. Lawn, R. Wilshaw, Review: indentation fracture: principles and applications. *J. Mater. Sci.* **10**, 1049–1081 (1975)
- J.W. Lee, I.K. Lloyd, H. Chai, Y.G. Jung, B.R. Lawn, Arrest, deflection, penetration and reinitiation of cracks in brittle layers across adhesive interlayers. *Acta Mater.* **55**, 5859–5866 (2007)
- W. Lili, Z. Xixiong, S. Shaoqiu, G. Su, and B. Hesheng, An impact dynamics investigation on some problems in bird strike on windshields of high speed aircrafts. *Acta Aeromuticaet Astronautica Sinica*, **12**(2), B27–B33 (1991)
- B.B. Mandelbrot, *Fractal geometry of nature*. WH Freeman and Company (1983)
- M.A. Martínez, I.S. Chocron, J. Rodríguez, V.S. Gálvez, L.A. Sastre, Confined compression of elastic adhesives at high rates of strain. *Int. J. Adhes. Adhes.* **18**, 375–383 (1998)
- B. Michel, L.B. Freund, *Dynamic Fracture Mechanics* (Cambridge University Press, Cambridge, 1990.) XVII, 563 pp., L 40.00 H/b. ISBN 0-521-30330-3 (Cambridge Monographs on Mechanics and Applied Mathematics), *Zeitschrift Angewandte Mathematik Und Mechanik*, **72** (1992) 383–384
- J.H. Nielsen, J.F. Olesen, H. Stang, The fracture process of tempered soda-lime-silica glass. *Exp. Mech.* **49**, 855–870 (2009)
- H. Park, W.W. Chen, Experimental investigation on dynamic crack propagating perpendicularly through interface in glass. *J Appl. Mech-t ASME* **78**(5) (2011)
- K. Ravi-Chandar, W.G. Knauss, An experimental investigation into dynamic fracture: IV. On the interaction of stress waves with propagating cracks. *Int. J. Fract.* **26**, 189–200 (1984)
- E. Sharon, J. Fineberg, Confirming the continuum theory of dynamic brittle fracture for fast cracks. *Nature* **397**, 333–335 (1999)
- E. Sharon, G. Cohen, J. Fineberg, Crack front waves and the dynamics of a rapidly moving crack. *Phys. Rev. Lett.* **88**, 47–103 (2002)
- Y.P. Shen, Y.H. Chen, Y.F. Pen, The finite element method of viscoelastic large deformation plane problem with Kirchhoff stress tensors-green strain tensors constitutive relation. *Acta Mech. Solida Sin.* 87–91 (1987)
- R.P. Singh, V. Parameswaran, An experimental investigation of dynamic crack propagation in a brittle material reinforced with a ductile layer. *Opt. Lasers Eng.* **40**, 289–306 (2003)
- J.H. Song, H. Wang, T. Belytschko, A comparative study on finite element methods for dynamic fracture. *Comput. Mech.* **42**, 239–250 (2008)
- M.G. Stout, D.A. Koss, C. Liu, J. Idasetima, Damage development in carbon/epoxy laminates under quasi-static and dynamic loading. *Compos. Sci. Technol.* **59**, 2339–2350 (1999)
- X. Sun, M.A. Khaleel, X. Sun, M.A. Khaleel, Effects of different design parameters on stone-impact resistance of automotive windshields. *Proc. Inst. Mech. Eng. Part D J. Automob. Eng.* **219**, 1059–1067 (2005)
- X. Sun, W. Liu, W. Chen, D. Templeton, Modeling and characterization of dynamic failure of borosilicate glass under compression/shear loading ☆. *Int. J. Impact Eng.* **36**, 226–234 (2009)
- S.R. Swanson, L.W. Christensen, A constitutive formulation for high-elongation propellants. *J. Spacecr. Rocket.* **20**, 559–566 (2015)
- A. Thom, Experimental and modeling studies of stress wave propagation in multilayer composite materials: low modulus interlayer effects. *J. Compos. Mater.* **39**, 981–1005 (2005)
- M. Timmel, S. Kolling, P. Osterrieder, P.A.D. Bois, A finite element model for impact simulation with laminated glass. *Int. J. Impact Eng.* **34**, 1465–1478 (2007)

- N. Vandenberghe, R. Vermorel, E. Villermaux, Star-shaped crack pattern of broken windows. *Phys. Rev. Lett.* **110**, 285–291 (2013)
- H.W. Wang, Z.J. Lai, X.U. Sun, D.J. Ming-Qiao, S.Q. Huang, Shi, Dynamic deformation and fracture of polymers taking account of damage evolution. *J Ningbo University*, **16**(4), 373–381 (2003)
- L.L. Wang, X.L. Dong, Z.J. Sun, Dynamic constitutive behavior of materials at high strain rate taking account of damage evolution. *Explosion Shock Waves* **26**, 193–198 (2006)
- X. Wang, Z. Feng, F. Wang, Z. Yue, Dynamic response analysis of bird strike on aircraft windshield based on damage-modified nonlinear viscoelastic constitutive relation. *Chinese J. Aeronaut.* **20**, 511–517 (2007)
- J. Xu, Y. Li, Model of vehicle velocity calculation in vehicle-pedestrian accident based on deflection of windshield. *J. Mech. Eng.* **45**, 210–215 (2009)
- J. Xu, Y. Li, G. Lu, W. Zhou, Reconstruction model of vehicle impact speed in pedestrian–vehicle accident. *Int. J. Impact Eng.* **36**, 783–788 (2009)
- J. Xu, Y. Li, X. Chen, Y. Yan, D. Ge, M. Zhu, B. Liu, Characteristics of windshield cracking upon low-speed impact: numerical simulation based on the extended finite element method. *Comput. Mater. Sci.* **48**, 582–588 (2010)
- J. Xu, Y. Li, B. Liu, M. Zhu, D. Ge, Experimental study on mechanical behavior of PVB laminated glass under quasi-static and dynamic loadings. *Compos. Part B* **42**, 302–308 (2011a)
- J. Xu, Y.B. Li, X. Chen, D.Y. Ge, B.H. Liu, M.Y. Zhu, T.H. Park, Automotive windshield – pedestrian head impact: energy absorption capability of interlayer material. *Int. J. Automot. Technol.* **12**, 687–695 (2011b)
- J. Xu, Y. Sun, B. Liu, M. Zhu, X. Yao, Y. Yan, Y. Li, X. Chen, Experimental and macroscopic investigation of dynamic crack patterns in PVB laminated glass sheets subject to light-weight impact. *Eng. Fail. Anal.* **18**, 1605–1612 (2011c)
- J. Xu, Y. Li, X. Chen, D. Ge, B. Liu, M. Zhu, T.H. Park, Automotive windshield – pedestrian head impact: energy absorption capability of interlayer material. *Int. J. Automot. Technol.* **12**, 687–695 (2011d)
- X.X. Zhang, R.C. Yu, G. Ruiz, M. Tarifa, M.A. Camara, Effect of loading rate on crack velocities in HSC. *Int. J. Impact Eng.* **37**, 359–370 (2010)
- S. Zhao, L.R. Dharani, X. Liang, L. Chai, S.D. Barbat, Crack initiation in laminated automotive glazing subjected to simulated head impact. *Int. J. Crashworthiness* **10**, 229–236 (2005)
- S.M. Zhao, L.R. Dharani, L. Chai, S.D. Barbat, Analysis of damage in laminated automotive glazing subjected to simulated head impact. *Eng. Fail. Anal.* **13**, 582–597 (2006a)
- S. Zhao, L.R. Dharani, L. Chai, S.D. Barbat, Dynamic response of laminated automotive glazing impacted by spherical featureless headform. *Int. J. Crashworthiness* **11**, 105–114 (2006b)
- F. Zhou, L. Wang, S. Hu, A damage-modified nonlinear VISCO-elastic constitutive relation and failure criterion of PMMA at high strain-rates. *Combust. Explosion Shock Waves* **12**, 333–342 (1992)

Part II
**Micromorphic and Cosserat in Gradient
Plasticity for Instabilities in Materials
and Structures**



Micromorphic Approach to Gradient Plasticity and Damage

15

Samuel Forest

Contents

Introduction	500
The Micromorphic Theory After Eringen and Mindlin and Its Extension to Plasticity	503
<i>Kinematics of Micromorphic Media</i>	503
<i>Principle of Virtual Power</i>	504
<i>Elastoviscoplasticity of Micromorphic Media</i>	507
The Micromorphic Approach to Various Gradient Field Theories	510
<i>Thermomechanics with Additional Degrees of Freedom</i>	510
<i>Non-dissipative Contribution of Generalized Stresses</i>	512
<i>Viscous Generalized Stress and Phase Field Model</i>	514
<i>Elastic-Plastic Decomposition of Generalized Strains</i>	516
Application to a Continuum Damage Model for Single Crystals and Its Regularization	517
<i>Constitutive Equations</i>	517
<i>Microdamage Continuum</i>	520
<i>Numerical Examples</i>	522
The Micromorphic Approach at Finite Deformations	523
<i>Micromorphic and Gradient Hyperelasticity</i>	524
<i>Finite Deformation Micromorphic Elastoviscoplasticity Using an Additive Decomposition of a Lagrangian Strain</i>	529
<i>Finite Deformation Micromorphic Viscoplasticity Using Local Objective Frames</i>	531
<i>Finite Deformation Micromorphic Elastoviscoplasticity Based on the Multiplicative Decomposition</i>	534
Conclusion	540
References	542

S. Forest (✉)

Centre des Materiaux, Mines ParisTech CNRS, PSL Research University, CNRS UMR 7633, Evry Cedex, Paris, France
e-mail: samuel.forest@mines-paristech.fr

© Springer Nature Switzerland AG 2019

G. Z. Voyiadjis (ed.), *Handbook of Nonlocal Continuum Mechanics for Materials and Structures*, https://doi.org/10.1007/978-3-319-58729-5_9

499

Abstract

Eringen and Mindlin’s original micromorphic continuum model is presented and extended towards finite elastic-plastic deformations. The framework is generalized to any additional kinematic degrees of freedom related to plasticity and/or damage mechanisms. It provides a systematic method to develop size-dependent plasticity and damage models, closely related to phase field approaches, that can be applied to hardening and/or softening material behavior. The regularization power of the method is illustrated in the case of damage in single crystals. Special attention is given to the various possible finite deformation formulations enhancing existing frameworks for finite elastoplasticity and damage.

Keywords

Gradient plasticity · Gradient damage · Micromorphic media · Regularization · Finite deformations · Generalized continua · Microstrain · Microstretch · Strain localization · Elasto-plasticity · Cleavage

Introduction

Micromorphic media are examples of three-dimensional generalized continua including additional degrees of freedom complementing the usual displacement vector. A classification of generalised mechanical continuum theories is proposed in Fig. 1 in order to locate more precisely the class of micromorphic media. The present chapter is limited to continuum media fulfilling the principle of local action, meaning that the mechanical state at a material point \underline{X} depends on variables defined at this point only (Truesdell and Toupin 1960; Truesdell and Noll 1965). The classical Cauchy continuum is called *simple material* because its response at

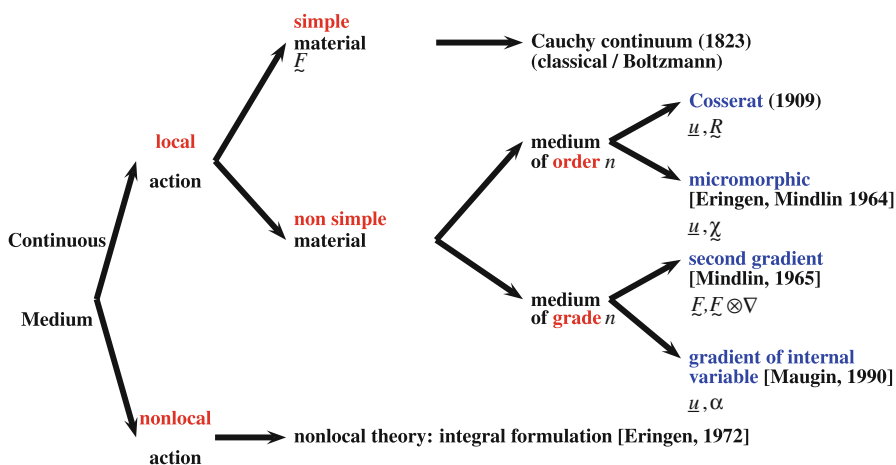


Fig. 1 A classification of the mechanics of generalised continua.

material point \underline{X} to deformations homogeneous in a neighborhood of \underline{X} determines uniquely its response to every deformation at \underline{X} . In *higher grade* materials, homogeneous deformations are not sufficient to characterise the material behaviour because they are sensitive to higher gradients of the displacement field. Mindlin formulated for instance the theories that include the second and third gradients of the displacement field (Mindlin 1965). The gradient effect may be limited to the plastic part of deformation which leads to strain gradient plasticity models (Aifantis 1984; Forest and Bertram 2011) or, more generally, theories that include the gradient of some internal variables (Maugin 1990). *Higher order* materials are characterised by additional degrees of freedom of the material points (Eringen 1999). Directors can be attached to each material point that evolve in a different way from the material lines. Cosserat directors can rotate. In the micromorphic continuum designed by Eringen and Suhubi (1964) and Mindlin (1964), the directors can also be distorted, so that a second order tensor is attributed to each material point. Tensors of higher order can even be introduced as proposed in Germain's general micromorphic theory (Germain 1973b; Forest and Sab 2017).

Higher order media are sometimes called continua with *microstructure*. This name is misleading because Cauchy material models can also integrate some aspects of the underlying microstructure as illustrated by classical homogenisation methods used to derive the effective properties of composites. However generalised continua incorporate a feature of the microstructure which is not accounted for by standard homogenisation methods, namely their size-dependent material response. They involve intrinsic lengths directly stemming from the microstructure of the material. The mechanics of generalized continua represents a way of introducing, in the continuum description of materials, some characteristic length scales associated with their microstructure (Mühlhaus 1995). Such intrinsic lengths and generalized constitutive equations can be identified in two ways. Direct identification is possible from experimental curves exhibiting clear size effects in plasticity or fracture or from full-field strain measurements of strongly heterogeneous fields (Geers et al. 1998). The effective properties of such generalized continua can also be derived from scale transition and homogenization techniques by prescribing appropriate boundary conditions on a representative volume of material with microstructure (Caillaud et al. 2003).

The multiplication of generalized continuum model formulations from Cosserat to strain gradient plasticity in literature may leave an impression of disorder and inconsistency. Recent accounts have shown, on the contrary, that unifying presentations of several classes of generalized continuum theories are possible (Hirschberger and Steinmann 2009; Forest 2009). One of them, called the micromorphic approach, encompasses most theories incorporating additional degrees of freedom from the well-established Cosserat, microstretch and micromorphic continua (Eringen 1999) up to Aifantis and Gurtin strain gradient plasticity theories. Gradient theories are obtained from the micromorphic approach by imposing some internal constraints linking the additional degrees of freedom and other model variables.

The micromorphic theory now arouses strong interest from the materials science and computational mechanics communities because of its regularisation power in

the context of softening plasticity and damage and of its rather simple implementation in a finite element program. The number of degrees of freedom is not an obstacle any more with constantly increasing computer power and parallel solvers.

The objective of this chapter is first to present the elastoviscoplasticity theory of micromorphic media at finite deformation. This presentation is based on the fundamental work of Eringen and on recent developments in the context of plasticity. It represents an update of the corresponding chapter in the CISM book (Forest 2012). The second objective of this chapter is to present an extension of the micromorphic theory to other kinds of additional degrees of freedom like plasticity and damage related variables. This micromorphic approach to gradient effects in materials' behavior is a systematic method for incorporating intrinsic lengths in non-linear continuum mechanical models. It is illustrated here in the case of an anisotropic plasticity and damage model. The so-called *microdamage* model takes into account the crystallography of plasticity and fracture in metal single crystals.

The nonlinear theory of micromorphic media is presented in section “[The Micromorphic Theory After Eringen and Mindlin and Its Extension to Plasticity.](#)” The micromorphic approach is exposed in section “[The Micromorphic Approach to Various Gradient Field Theories](#)” together with the closely related phase field approach. Differences and similarities between the micromorphic framework and the phase field approach are pointed out following the general framework provided in Gurtin (1996). A single crystal plasticity and damage model is explored in section “[Application to a Continuum Damage Model for Single Crystals and Its Regularization](#)” up to crack propagation simulation. This presentation of the micromorphic approach and the corresponding example are taken from the formulation presented in Aslan and Forest (2011). A first account of the method is given at small strains for the sake of simplicity. It is followed by extensions to finite deformations, following the guidelines given in Forest (2016). The presentation is limited to the static case, the reader being referred to Eringen's original dynamical formulations and to Forest and Sab (2017) for the consideration of inertial terms.

Intrinsic notations are used throughout the chapter. In particular, scalars, vectors, tensors of second, third and fourth ranks are denoted by a , \underline{a} , $\underline{\underline{a}}$, $\underline{\underline{\underline{a}}}$, respectively. Contractions are written as:

$$\underline{\underline{a}} : \underline{\underline{b}} = a_{ij} b_{ij}, \underline{\underline{\underline{a}}} : \underline{\underline{\underline{b}}} = a_{ijk} b_{ijk}, \underline{\underline{\underline{\underline{a}}}} :: \underline{\underline{\underline{\underline{b}}}} = a_{ijkl} b_{ijkl} \quad (1)$$

using the Einstein summation rule for repeated indices. The tensor product is denoted by \otimes . For example, the component $(\underline{\underline{a}} \otimes \underline{\underline{b}})_{ijkl}$ is $a_{ij} b_{kl}$. A modified tensor product \boxtimes is also used: the component $(\underline{\underline{a}} \boxtimes \underline{\underline{b}})_{ijkl}$ is $a_{ik} b_{jl}$.

The gradient operators ∇_x or ∇_X are introduced when the functions depend on microscopic coordinates \underline{x} or macroscopic coordinates \underline{X} . The following notation is used:

$$\underline{U} \otimes \nabla_X = U_{i,j} \underline{e}_i \otimes \underline{e}_j, \quad \text{with} \quad U_{i,j} = \frac{\partial U_i}{\partial X_j} \quad (2)$$

$$\underline{u} \otimes \nabla_x = u_{i,j} \underline{e}_i \otimes \underline{e}_j, \quad \text{with} \quad u_{i,j} = \frac{\partial u_i}{\partial x_j} \quad (3)$$

where $(\underline{e}_i)_{i=1,2,3}$ is a Cartesian orthonormal basis.

The Micromorphic Theory After Eringen and Mindlin and Its Extension to Plasticity

Kinematics of Micromorphic Media

The degrees of freedom of the theory are the displacement vector \underline{u} and the microdeformation tensor $\underline{\chi}$:

$$DOF := \left\{ \underline{u}, \underline{\chi} \right\}$$

The current position of the material point is given by the transformation Φ according to $\underline{x} = \Phi(\underline{X}, t) = \underline{X} + \underline{u}(\underline{X}, t)$. The microdeformation describes the deformation of a triad of directors, $\underline{\Xi}^i$ attached to the material point.

$$\underline{\xi}^i(\underline{X}, t) = \underline{\chi}(\underline{X}, t) \cdot \underline{\Xi}^i \quad (4)$$

As such, its determinant is taken as strictly positive. The polar decomposition of the generally incompatible microdeformation field $\underline{\chi}(\underline{X})$ is introduced

$$\underline{\chi} = \underline{R}^\sharp \cdot \underline{U}^\sharp \quad (5)$$

Internal constraints can be prescribed to the microdeformation. The micromorphic medium reduces to the Cosserat medium when the microdeformation is constrained to be a pure rotation: $\underline{\chi} \equiv \underline{R}^\sharp$. The microstrain medium is obtained when $\underline{\chi} \equiv \underline{U}^\sharp$ (Forest and Sievert 2006). Finally, the second gradient theory is retrieved when the microdeformation coincides with the deformation gradient, $\underline{\chi} \equiv \underline{F}$. A hierarchy of higher order continua can be established by specialising the micromorphic theory and depending on the targeted material class, see Table 1.

The following kinematical quantities are then introduced:

- the velocity field $\underline{v}(\underline{x}, t) := \dot{\underline{u}}(\Phi^{-1}(\underline{x}, t))$
- the deformation gradient $\underline{F} = \underline{1} + \underline{u} \otimes \nabla_x$

Table 1 A hierarchy of higher order continua

Name	Number of DOF	DOF	References
Cauchy	3	$\underline{\mathbf{u}}$	Cauchy (1822)
Microdilatation	4	$\underline{\mathbf{u}}, \chi$	Goodman and Cowin (1972) and Steeb and Diebels (2003)
Cosserat	6	$\underline{\mathbf{u}}, \underline{\mathbf{R}}$	Kafadar and Eringen (1971)
Microstretch	7	$\underline{\mathbf{u}}, \chi, \underline{\mathbf{R}}$	Eringen (1990)
Microstrain	9	$\underline{\mathbf{u}}, \underline{\mathbf{C}}^\sharp$	Forest and Sievert (2006)
Micromorphic	12	$\underline{\mathbf{u}}, \underline{\mathbf{\chi}}$	Eringen and Suhubi (1964) and Mindlin (1964)

- the velocity gradient $\underline{\mathbf{v}} \otimes \nabla_x = \underline{\dot{\mathbf{F}}} \cdot \underline{\mathbf{F}}^{-1}$
- the microdeformation rate $\underline{\dot{\chi}} \cdot \underline{\chi}^{-1}$
- the third rank Lagrangean microdeformation gradient $\underline{\mathbf{K}} := \underline{\chi}^{-1} \cdot \underline{\chi} \otimes \nabla_X$
- the gradient of the microdeformation rate tensor

$$\left(\underline{\dot{\chi}} \cdot \underline{\chi}^{-1} \right) \otimes \nabla_x = \underline{\chi} \cdot \underline{\mathbf{K}} : \left(\underline{\chi}^{-1} \boxtimes \underline{\mathbf{F}}^{-1} \right) \quad (6)$$

and the corresponding index notation:

$$\left(\underline{\dot{\chi}}_{il} \chi_{lj}^{-1} \right)_{,k} = \chi_{ip} \dot{K}_{pqr} \chi_{qj}^{-1} F_{rk}^{-1}$$

Principle of Virtual Power

The method of virtual power is used to introduce the generalised stress tensors and the field and boundary equations they must satisfy (Germain 1973b).

The modelling variables are introduced according to a first gradient theory:

$$MODEL = \left\{ \underline{\mathbf{v}}, \underline{\mathbf{v}} \otimes \nabla_x, \underline{\dot{\chi}} \cdot \underline{\chi}^{-1}, \left(\underline{\dot{\chi}} \cdot \underline{\chi}^{-1} \right) \otimes \nabla_x \right\}$$

The virtual power of internal forces of a subdomain $\mathcal{D} \subset \mathcal{B}$ of the body is

$$\mathcal{P}^{(i)} \left(\underline{\mathbf{v}}^*, \underline{\dot{\chi}}^* \cdot \underline{\chi}^{*-1} \right) = \int_{\mathcal{D}} p^{(i)} \left(\underline{\mathbf{v}}^*, \underline{\dot{\chi}}^* \cdot \underline{\chi}^{*-1} \right) dV$$

The virtual power density of internal forces is a linear form on the fields of virtual modeling variables:

$$\begin{aligned}
p^{(i)} &= \underline{\underline{\sigma}} : (\underline{\underline{\dot{F}}} \cdot \underline{\underline{F}}^{-1}) + \underline{\underline{s}} : (\underline{\underline{\dot{F}}} \cdot \underline{\underline{F}}^{-1} - \underline{\underline{\dot{\chi}}} \cdot \underline{\underline{\chi}}^{-1}) + \underline{\underline{M}} : \left((\underline{\underline{\dot{\chi}}} \cdot \underline{\underline{\chi}}^{-1}) \otimes \nabla_x \right) \\
&= \underline{\underline{\sigma}} : (\underline{\underline{\dot{F}}} \cdot \underline{\underline{F}}^{-1}) + \underline{\underline{s}} : \left(\underline{\underline{\chi}} \cdot (\underline{\underline{\chi}}^{-1} \cdot \underline{\underline{F}}) \cdot \underline{\underline{F}}^{-1} \right) \\
&\quad + \underline{\underline{M}} : \left(\underline{\underline{\chi}} \cdot \underline{\underline{\dot{\chi}}} : (\underline{\underline{\chi}}^{-1} \boxtimes \underline{\underline{F}}^{-1}) \right)
\end{aligned} \tag{7}$$

where the relative deformation rate $\underline{\underline{\dot{F}}} \cdot \underline{\underline{F}}^{-1} - \underline{\underline{\dot{\chi}}} \cdot \underline{\underline{\chi}}^{-1}$ is introduced and expressed in terms of the rate of the relative deformation $\underline{\underline{\chi}}^{-1} \cdot \underline{\underline{F}}$. The virtual power density of internal forces is invariant with respect to virtual rigid body motions so that $\underline{\underline{\sigma}}$ must be symmetric. The generalised stress tensors conjugate to the velocity gradient, the relative deformation rate and the gradient of the microdeformation rate are the simple stress tensor $\underline{\underline{\sigma}}$, the relative stress tensor $\underline{\underline{s}}$ and the double stress tensor $\underline{\underline{M}}$ of third rank.

The Gauss theorem is then applied to the power of internal forces.

$$\begin{aligned}
\int_{\mathcal{D}} p^{(i)} dV &= \int_{\partial \mathcal{D}} \underline{\underline{v}}^* \cdot (\underline{\underline{\sigma}} + \underline{\underline{s}}) \cdot \underline{\underline{n}} dS + \int_{\partial \mathcal{D}} (\underline{\underline{\dot{\chi}}}^* \cdot \underline{\underline{\chi}}^{*-1}) : \underline{\underline{M}} \cdot \underline{\underline{n}} dS \\
&\quad - \int_{\mathcal{D}} \underline{\underline{v}}^* \cdot (\underline{\underline{\sigma}} + \underline{\underline{s}}) \cdot \nabla_x dV \\
&\quad - \int_{\mathcal{D}} (\underline{\underline{\dot{\chi}}}^* \cdot \underline{\underline{\chi}}^{*-1}) : (\underline{\underline{M}} \cdot \nabla_x + \underline{\underline{s}}) \cdot dV
\end{aligned}$$

The form of the previous boundary integral dictates the possible form of the power of contact forces acting on the boundary $\partial \mathcal{D}$ of the subdomain $\mathcal{D} \subset \mathcal{B}$.

$$\begin{aligned}
\mathcal{P}^{(c)} \left(\underline{\underline{v}}^*, \underline{\underline{\dot{\chi}}}^* \cdot \underline{\underline{\chi}}^{*-1} \right) &= \int_{\partial \mathcal{D}} p^{(c)} \left(\underline{\underline{v}}^*, \underline{\underline{\dot{\chi}}}^* \cdot \underline{\underline{\chi}}^{*-1} \right) dV \\
p^{(c)} \left(\underline{\underline{v}}^*, \underline{\underline{\dot{\chi}}}^* \cdot \underline{\underline{\chi}}^{*-1} \right) &= \underline{\underline{t}} \cdot \underline{\underline{v}}^* + \underline{\underline{m}} : \left(\underline{\underline{\dot{\chi}}}^* \cdot \underline{\underline{\chi}}^{*-1} \right)
\end{aligned}$$

where the simple traction $\underline{\underline{t}}$ and double traction $\underline{\underline{m}}$ are introduced.

The power of forces acting at a distance is defined as

$$\begin{aligned}
\mathcal{P}^{(e)} \left(\underline{\underline{v}}^*, \underline{\underline{\dot{\chi}}}^* \cdot \underline{\underline{\chi}}^{*-1} \right) &= \int_{\mathcal{D}} p^{(e)} \left(\underline{\underline{v}}^*, \underline{\underline{\dot{\chi}}}^* \cdot \underline{\underline{\chi}}^{*-1} \right) dV \\
p^{(e)} \left(\underline{\underline{v}}^*, \underline{\underline{\dot{\chi}}}^* \cdot \underline{\underline{\chi}}^{*-1} \right) &= \underline{\underline{f}} \cdot \underline{\underline{v}}^* + \underline{\underline{p}} : \left(\underline{\underline{\dot{\chi}}}^* \cdot \underline{\underline{\chi}}^{*-1} \right)
\end{aligned}$$

including simple body forces \underline{f} and double body forces \underline{p} . More general double and triple volume forces could also be incorporated according to Germain (1973b).

The principle of virtual power is now stated in the static case,

$$\forall \underline{v}^*, \forall \underline{\chi}^*, \forall \mathcal{D} \subset \mathcal{B}, \mathcal{P}^{(i)} \left(\underline{v}^*, \underline{\dot{\chi}}^* \cdot \underline{\chi}^{*-1} \right) = \mathcal{P}^{(c)} \left(\underline{v}^*, \underline{\dot{\chi}}^* \cdot \underline{\chi}^{*-1} \right) + \mathcal{P}^{(e)} \left(\underline{v}^*, \underline{\dot{\chi}}^* \cdot \underline{\chi}^{*-1} \right)$$

This variational formulation leads to.

$$\begin{aligned} & \int_{\partial \mathcal{D}} \underline{v}^* \cdot (\underline{\sigma} + \underline{s}) \cdot \underline{n} \, ds + \int_{\partial \mathcal{D}} \left(\underline{\dot{\chi}}^* \cdot \underline{\chi}^{*-1} \right) : \underline{\underline{M}} \cdot \underline{n} \, dS \\ & - \int_{\mathcal{D}} \underline{v}^* \cdot \left((\underline{\sigma} + \underline{s}) \cdot \nabla_x + \underline{f} \right) \, dV \\ & - \int_{\mathcal{D}} \left(\underline{\dot{\chi}}^* \cdot \underline{\chi}^{*-1} \right) : \left(\underline{\underline{M}} \cdot \nabla_x + \underline{s} + \underline{p} \right) \, dV = 0 \end{aligned}$$

which delivers the field equations of the problem (Kirchner and Steinmann 2005; Lazar and Maugin 2007; Hirschberger et al. 2007):

- balance of momentum equation (static case)

$$\left(\underline{\sigma} + \underline{s} \right) \cdot \nabla_x + \underline{f} = 0, \quad \forall \underline{x} \in \mathcal{B} \quad (8)$$

- balance of generalized moment of momentum equation (static case)

$$\underline{\underline{M}} \cdot \nabla_x + \underline{s} + \underline{p} = 0, \quad \forall \underline{x} \in \mathcal{B} \quad (9)$$

- boundary conditions

$$\left(\underline{\sigma} + \underline{s} \right) \cdot \underline{n} = \underline{t}, \quad \forall \underline{x} \in \partial \mathcal{B} \quad (10)$$

$$\underline{\underline{M}} \cdot \underline{n} = \underline{m}, \quad \forall \underline{x} \in \partial \mathcal{B} \quad (11)$$

Elastoviscoplasticity of Micromorphic Media

This section is dedicated to the formulation of constitutive equations for micromorphic media. The general case of hyperelastic-viscoplastic materials is considered.

Elastic–Plastic Decomposition of the Generalised Strain Measures

According to Eringen (1999), the following Lagrangean strain measures are adopted:

$$STRAIN = \left\{ \underset{\sim}{C} := \underset{\sim}{F}^T \cdot \underset{\sim}{F}, \underset{\sim}{Y} := \underset{\sim}{\chi}^{-1} \cdot \underset{\sim}{F}, \underset{\sim}{K} := \underset{\sim}{\chi}^{-1} \cdot \left(\underset{\sim}{\chi} \otimes \nabla_X \right) \right\}$$

i.e. the Cauchy–Green strain tensor, the relative deformation and the microdeformation gradient.

In the presence of plastic deformation, the question arises of splitting the previous Lagrangean strain measures into elastic and plastic contributions. Following Mandel (1973), a multiplicative decomposition of the deformation gradient is postulated:

$$\underset{\sim}{F} = \underset{\sim}{F}^e \cdot \underset{\sim}{F}^p = \underset{\sim}{R}^e \cdot \underset{\sim}{U}^e \cdot \underset{\sim}{F}^p \tag{12}$$

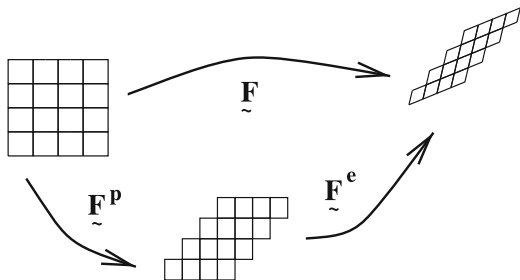
which defines an intermediate local configuration at each material point, see Fig. 2. Uniqueness of the decomposition requires the suitable definition of directors. Such directors are available in any micromorphic theory.

A multiplicative decomposition of the microdeformation is also considered:

$$\underset{\sim}{\chi} = \underset{\sim}{\chi}^e \cdot \underset{\sim}{\chi}^p = \underset{\sim}{R}^{e\sharp} \cdot \underset{\sim}{U}^{e\sharp} \cdot \underset{\sim}{\chi}^p \tag{13}$$

according to Forest and Sievert (2003, 2006). The uniqueness of the decomposition also requires the suitable definition of directors. As an example, lattice directions in a single crystal are physically relevant directors for an elastoviscoplasticity micromorphic theory, see Aslan et al. (2011). Finally, a partition rule must also be proposed for the third strain measure, namely the microdeformation gradient. Sansour (1998a, b) introduced an additive decomposition of curvature:

Fig. 2 Multiplicative decomposition of the deformation gradient



$$\underline{\underline{\mathbf{K}}} = \underline{\underline{\mathbf{K}}}^e + \underline{\underline{\mathbf{K}}}^p \quad (14)$$

A quasi-additive decomposition was proposed by Forest and Sievert (2003) with the objective of defining an intermediate local configuration for which all generalised stress tensor are simultaneously released, as it will become apparent in the next section:

$$\underline{\underline{\mathbf{K}}} = \underline{\underline{\chi}}^{p-1} \cdot \underline{\underline{\mathbf{K}}}^e : \left(\underline{\underline{\chi}}^p \boxtimes \underline{\underline{\mathbf{F}}}^p \right) + \underline{\underline{\mathbf{K}}}^p \quad (15)$$

Constitutive Equations

The continuum thermodynamic formulation is essentially unchanged in the presence of additional degrees of freedom provided that all functionals are properly extended to the new sets of variables. The local equation of energy balance is written in its usual form:

$$\rho \dot{\varepsilon} = p^{(i)} - \underline{\mathbf{q}} \cdot \nabla + r \quad (16)$$

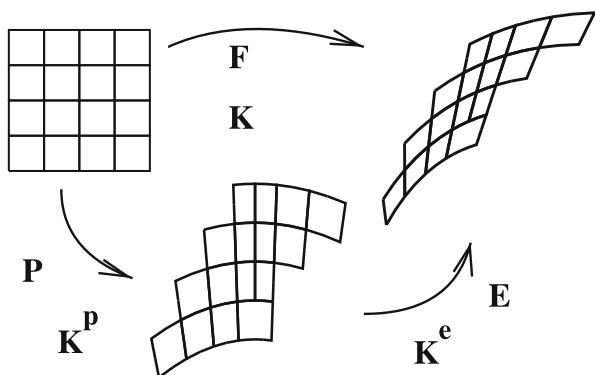
where ε is the specific internal energy density, and $p^{(i)}$ is the power density of internal forces according to Eq. (7). The heat flux vector is $\underline{\mathbf{q}}$ and r is a heat source term (Fig. 3). The local form of the second principle of thermodynamics is written as.

$$\rho \dot{\eta} + \left(\frac{\underline{\mathbf{q}}}{T} \right) \cdot \nabla - \frac{r}{T} \geq 0$$

where η is the specific entropy density. Introducing the Helmholtz free energy function the ψ , second law becomes

$$p^{(i)} - \rho \dot{\psi} - \eta \dot{T} - \frac{\underline{\mathbf{q}}}{T} \cdot (\nabla T) \geq 0$$

Fig. 3 Definition of an intermediate local configuration for micromorphic elastoplasticity



The state variables of the elastoviscoplastic micromorphic material are all the elastic strain measures and a set of internal variables q . The free energy density is a function of the state variables:

$$\Psi \left(\underline{\underline{\mathbf{C}}}^e := \underline{\underline{\mathbf{F}}}^{eT} \cdot \underline{\underline{\mathbf{F}}}^e, \underline{\underline{\boldsymbol{\Upsilon}}}^e := \underline{\underline{\boldsymbol{\chi}}}^{e-1} \cdot \underline{\underline{\mathbf{F}}}^e, \underline{\underline{\mathbf{K}}}^e, q \right)$$

The exploitation of the entropy inequality leads to the definition of the hyperelastic state laws in the form:

$$\begin{aligned} \underline{\underline{\boldsymbol{\sigma}}} &= 2\underline{\underline{\mathbf{F}}}^e \cdot \rho \frac{\partial \Psi}{\partial \underline{\underline{\mathbf{C}}}^e} \cdot \underline{\underline{\mathbf{F}}}^{eT}, \underline{\underline{\mathbf{s}}} = \underline{\underline{\mathbf{R}}}^{e\sharp} \cdot \underline{\underline{\mathbf{U}}}^{e\sharp-1} \cdot \rho \frac{\partial \Psi}{\partial \underline{\underline{\boldsymbol{\Upsilon}}}^e} \cdot \underline{\underline{\mathbf{F}}}^{eT} \\ \underline{\underline{\mathbf{M}}} &= \underline{\underline{\boldsymbol{\chi}}}^{-T} \cdot \rho \frac{\partial \Psi}{\partial \underline{\underline{\mathbf{K}}}^e} : \left(\underline{\underline{\boldsymbol{\chi}}}^T \boxtimes \underline{\underline{\mathbf{F}}}^T \right) \end{aligned} \quad (17)$$

while the entropy density is still given by $\eta = -\frac{\partial \Psi}{\partial T}$. The thermodynamic force associated with the internal variable q is

$$R = -\rho \frac{\partial \Psi}{\partial q}$$

The hyperelasticity law (17) for the double stress tensor was derived for the additive decomposition (14). The quasi-additive decomposition (15) leads to an hyperelastic constitutive equation for the conjugate stress $\underline{\underline{\mathbf{M}}}$ in the current configuration, that has also the same form as for pure hyperelastic behaviour. One finds:

$$\underline{\underline{\mathbf{M}}} = \underline{\underline{\boldsymbol{\chi}}}^{e-T} \cdot \rho \frac{\partial \Psi}{\partial \underline{\underline{\mathbf{K}}}^e} : \left(\underline{\underline{\boldsymbol{\chi}}}^{eT} \boxtimes \underline{\underline{\mathbf{F}}}^{eT} \right) \quad (18)$$

The residual intrinsic dissipation is

$$D = \underline{\underline{\Sigma}} : \left(\underline{\underline{\dot{\mathbf{F}}}}^p \cdot \underline{\underline{\mathbf{F}}}^{p-1} \right) + \underline{\underline{\mathcal{S}}} : \left(\underline{\underline{\dot{\boldsymbol{\chi}}}}^p \cdot \underline{\underline{\boldsymbol{\chi}}}^{p-1} \right) + \underline{\underline{\mathcal{M}}} : \underline{\underline{\dot{\mathbf{K}}}}^p + R\dot{q} \geq 0$$

where generalised Mandel stress tensors have been defined.

$$\underline{\underline{\Sigma}} = \underline{\underline{\mathbf{F}}}^{eT} \cdot \left(\underline{\underline{\boldsymbol{\sigma}}} + \underline{\underline{\mathbf{s}}} \right) \cdot \underline{\underline{\mathbf{F}}}^{e-T}, \underline{\underline{\mathcal{S}}} = -\underline{\underline{\mathbf{U}}}^{e\sharp} \cdot \underline{\underline{\mathbf{R}}}^{e\sharp T} \cdot \underline{\underline{\mathbf{s}}} \cdot \underline{\underline{\mathbf{R}}}^{e\sharp} \cdot \underline{\underline{\mathbf{U}}}^{e\sharp-1} \quad (19)$$

$$\underline{\underline{\mathcal{M}}} = \underline{\underline{\boldsymbol{\chi}}}^T \cdot \underline{\underline{\mathcal{S}}} : \left(\underline{\underline{\boldsymbol{\chi}}}^{-T} \boxtimes \underline{\underline{\mathbf{F}}}^{-T} \right) \quad (20)$$

At this stage, one may define a dissipation potential, function of the Mandel stress tensors, from which the viscoplastic flow rule and the evolution equations for the internal variables are derived.

$$\dot{\tilde{\mathbf{F}}}^p \cdot \tilde{\mathbf{F}}^{p-1} = \frac{\partial \Omega}{\partial \tilde{\boldsymbol{\Sigma}}}, \dot{\tilde{\boldsymbol{\chi}}}^p \cdot \tilde{\boldsymbol{\chi}}^{p-1} = \frac{\partial \Omega}{\partial \tilde{\boldsymbol{\mathcal{S}}}}, \dot{\tilde{\mathbf{K}}}^p = \frac{\partial \Omega}{\partial \tilde{\boldsymbol{\mathcal{M}}}}, \dot{q} = \frac{\partial \Omega}{\partial R}$$

The convexity of the dissipation potential with respect to its arguments ensures the positivity of the dissipation rate at each instant.

Explicit constitutive equations can be found in Forest and Sievert (2003), Grammenoudis and Tsakmakis (2009), Grammenoudis et al. (2009), Regueiro (2010), and Sansour et al. (2010). Examples of application of elastoplastic micromorphic media can be found in Dillard et al. (2006) for plasticity and failure of metallic foams.

The Micromorphic Approach to Various Gradient Field Theories

The previous micromorphic model can be extended to other types of additional degrees of freedom. This leads to a systematic approach for the construction of generalized continuum models with enriched kinematics. The method is presented first within the small deformation framework. The general formulation is discussed in the next section.

Thermomechanics with Additional Degrees of Freedom

One starts from an elastoviscoplasticity model formulation within the framework of the classical Cauchy continuum and classical continuum thermodynamics according to Germain et al. (1983) and Maugin (1999). The material behaviour is characterized by the reference sets of degrees of freedom and state variables.

$$DOF0 = \{\underline{\mathbf{u}}\}, STATE0 = \{\underline{\boldsymbol{\varepsilon}}, T, q\} \quad (21)$$

which the free energy density function ψ may depend on. The small strain tensor is denoted by $\underline{\boldsymbol{\varepsilon}}$ whereas q represents the whole set of internal variables of arbitrary tensorial order accounting for nonlinear processes at work inside the material volume element, like isotropic and kinematic hardening variables. The absolute temperature is T .

Additional degrees of freedom ϕ_χ are then introduced in the previous original model. They may be of any tensorial order and of different physical nature (deformation, plasticity or damage variable). The notation χ indicates that these variables eventually represent some microstructural features of the material so that we will call them micromorphic variables or microvariables (*microdeformation, microdamage...*). The *DOF* and *STATE* spaces are extended as follows:

$$DOF = \{\underline{\mathbf{u}}, \phi_\chi\}, STATE = \{\underline{\boldsymbol{\varepsilon}}, T, q, \phi_\chi, \nabla\phi_\chi\} \quad (22)$$

Depending on the physical nature of ϕ_χ , it may or may not be a state variable. For instance, if the microvariable is a microrotation as in the Cosserat model, it is not a state variable for objectivity reasons and will appear in *STATE* only in combination with the macrorotation. In contrast, if the microvariable is a microplastic equivalent strain, as in Aifantis model, it then explicitly appears in the state space.

The virtual power of internal forces is then extended to the power done by the micromorphic variable and its first gradient:

$$\begin{aligned} \mathcal{P}^{(i)}(\underline{\mathbf{v}}^*, \dot{\phi}_\chi^*) &= - \int_{\mathcal{D}} p^{(i)}(\underline{\mathbf{v}}^*, \dot{\phi}_\chi^*) dV \\ p^{(i)}(\underline{\mathbf{v}}^*, \dot{\phi}_\chi^*) &= \underline{\boldsymbol{\sigma}} : \nabla \underline{\mathbf{v}}^* + a \dot{\phi}_\chi^* + \underline{\mathbf{b}} \cdot \nabla \dot{\phi}_\chi^* \end{aligned} \quad (23)$$

where \mathcal{D} is a subdomain of the current configuration of the body. The Cauchy stress is $\underline{\boldsymbol{\sigma}}$ and a and $\underline{\mathbf{b}}$ are generalized stresses associated with the micromorphic variable and its first gradient. Similarly, the power of contact forces must be extended as follows:

$$\mathcal{P}^{(c)}(\underline{\mathbf{v}}^*, \dot{\phi}_\chi^*) = \int_{\mathcal{D}} p^{(c)}(\underline{\mathbf{v}}^*, \dot{\phi}_\chi^*) dV, \quad p^{(c)}(\underline{\mathbf{v}}^*, \dot{\phi}_\chi^*) = \underline{\mathbf{t}} \cdot \underline{\mathbf{v}}^* + a^c \dot{\phi}_\chi^* \quad (24)$$

where $\underline{\mathbf{t}}$ is the traction vector and a^c a generalized traction. For conciseness, we do not extend the power of forces acting at a distance and keep the classical form:

$$\mathcal{P}^{(e)}(\underline{\mathbf{v}}^*, \dot{\phi}_\chi^*) = \int_{\mathcal{D}} p^{(e)}(\underline{\mathbf{v}}^*, \dot{\phi}_\chi^*) dV, \quad p^{(e)}(\underline{\mathbf{v}}^*, \dot{\phi}_\chi^*) = \rho \underline{\mathbf{f}} \cdot \underline{\mathbf{v}}^* \quad (25)$$

where $\rho \underline{\mathbf{f}}$ accounts for given simple body forces. Following Germain (1973a), given body couples and double forces working with the gradient of the velocity field could also be introduced in the theory. The generalized principle of virtual power with respect to the velocity and micromorphic variable fields, is presented here in the static case only:

$$\mathcal{P}^{(i)}(\underline{\mathbf{v}}^*, \dot{\phi}_\chi^*) + \mathcal{P}^{(e)}(\underline{\mathbf{v}}^*, \dot{\phi}_\chi^*) + \mathcal{P}^{(c)}(\underline{\mathbf{v}}^*, \dot{\phi}_\chi^*) = 0, \quad \forall \mathcal{D} \subset \Omega, \quad \forall \underline{\mathbf{v}}^*, \dot{\phi}_\chi^* \quad (26)$$

The method of virtual power according to Maugin (1980) is used then to derive the standard local balance of momentum equation:

$$\operatorname{div} \underline{\boldsymbol{\sigma}} + \rho \underline{\mathbf{f}} = 0, \quad \forall \underline{\mathbf{x}} \in \Omega \quad (27)$$

and the generalized balance of micromorphic momentum equation:

$$\operatorname{div} \underline{\mathbf{b}} - a = 0, \quad \forall \underline{\mathbf{x}} \in \Omega \quad (28)$$

The method also delivers the associated boundary conditions for the simple and generalized tractions:

$$\underline{t} = \underline{\sigma} \cdot \underline{n}, a^c = \underline{b} \cdot \underline{n}, \forall \underline{x} \in \partial \mathcal{D} \quad (29)$$

The local balance of energy is also enhanced by the generalized micromorphic power already included in the power of internal forces (23):

$$\rho \dot{\epsilon} = p^{(i)} - \operatorname{div} \underline{q} + \rho r \quad (30)$$

where ϵ is the specific internal energy, \underline{q} the heat flux vector and r denotes external heat sources. The entropy principle takes the usual local form:

$$-\rho (\dot{\psi} + \eta \dot{T}) + p^{(i)} - \frac{\underline{q}}{T} \cdot \nabla T \geq 0 \quad (31)$$

where it is assumed that the entropy production vector is still equal to the heat vector divided by temperature, as in classical thermomechanics according to Coleman and Noll (1963). Again, the enhancement of the theory goes through the enriched power density of internal forces (7). The entropy principle is exploited according to classical continuum thermodynamics to derive the state laws. At this stage it is necessary to be more specific on the dependence of the state functions ψ , η , $\underline{\sigma}$, a , \underline{b} on state variables and to distinguish between dissipative and non-dissipative mechanisms. The introduction of dissipative mechanisms may require an increase in the number of state variables. These different situations are considered in the following subsections.

Non-dissipative Contribution of Generalized Stresses

Dissipative events are assumed here to enter the model only via the classical mechanical part. Total strain is split into elastic and plastic parts:

$$\underline{\epsilon} = \underline{\tilde{\epsilon}}^e + \underline{\tilde{\epsilon}}^p \quad (32)$$

The following constitutive functional dependencies are then introduced

$$\begin{aligned} \psi &= \hat{\psi}(\underline{\tilde{\epsilon}}^e, T, q, \phi_\chi, \nabla \phi_\chi), \quad \underline{\sigma} = \hat{\underline{\sigma}}(\underline{\tilde{\epsilon}}^e, T, q, \phi_\chi, \nabla \phi_\chi), \\ \eta &= \hat{\eta}(\underline{\tilde{\epsilon}}^e, T, q, \phi_\chi, \nabla \phi_\chi), \quad a = \hat{a}(\underline{\tilde{\epsilon}}^e, T, q, \phi_\chi, \nabla \phi_\chi), \\ \underline{b} &= \hat{\underline{b}}(\underline{\tilde{\epsilon}}^e, T, q, \phi_\chi, \nabla \phi_\chi) \end{aligned} \quad (33)$$

The entropy inequality (31) can be expanded as:

$$\begin{aligned} & \left(\underline{\underline{\sigma}} - \rho \frac{\partial \widehat{\psi}}{\partial \underline{\underline{\boldsymbol{\varepsilon}}^e}} \right) : \dot{\underline{\underline{\boldsymbol{\varepsilon}}^e} } + \rho \left(\eta + \frac{\partial \widehat{\psi}}{\partial T} \right) \dot{T} + \left(a - \rho \frac{\partial \widehat{\psi}}{\partial \phi_\chi} \right) \dot{\phi}_\chi \\ & + \left(\underline{\underline{\mathbf{b}}} - \rho \frac{\partial \widehat{\psi}}{\partial \nabla \phi_\chi} \right) \cdot \nabla \dot{\phi}_\chi + \underline{\underline{\boldsymbol{\sigma}}} : \dot{\underline{\underline{\boldsymbol{\varepsilon}}^p} } - \rho \frac{\partial \widehat{\psi}}{\partial q} \dot{q} - \frac{q}{T} \cdot \nabla T \geq 0 \end{aligned} \quad (34)$$

Assuming that no dissipation is associated with the four first terms of the previous inequality, the following state laws are found.

$$\underline{\underline{\boldsymbol{\sigma}}} = \rho \frac{\partial \widehat{\psi}}{\partial \underline{\underline{\boldsymbol{\varepsilon}}^e}}, \quad \eta = -\frac{\partial \widehat{\psi}}{\partial T}, \quad R = -\rho \frac{\partial \widehat{\psi}}{\partial q} \quad (35)$$

$$a = \rho \frac{\partial \widehat{\psi}}{\partial \phi_\chi}, \quad \underline{\underline{\mathbf{b}}} = \rho \frac{\partial \widehat{\psi}}{\partial \nabla \phi_\chi} \quad (36)$$

and the residual dissipation is.

$$D^{res} = W^p + R\dot{q} - \frac{q}{T} \cdot \nabla T \geq 0 \quad (37)$$

where W^p represents the (visco-)plastic power and R the thermodynamic force associated with the internal variable q . The existence of a convex dissipation potential, $\Omega(\underline{\underline{\boldsymbol{\sigma}}}, R)$ depending on the thermodynamic forces can then be assumed from which the evolution rules for internal variables are derived, that identically fulfill the entropy inequality, as usually done in classical continuum thermomechanics (Germain et al. 1983):

$$\dot{\underline{\underline{\boldsymbol{\varepsilon}}^p} } = \frac{\partial \Omega}{\partial \underline{\underline{\boldsymbol{\sigma}}}}, \quad \dot{q} = \frac{\partial \Omega}{\partial R} \quad (38)$$

Micromorphic Model

After presenting the general approach, we readily give the most simple example which provides a direct connection to several existing generalized continuum models. An element ϕ is selected in the STATE0 set, see Eq. (21) or among other variables present in the original model. Cases are first considered where ϕ and ϕ_χ are observer invariant quantities. The free energy density function ψ is chosen as a function of the generalized relative strain variable e defined as:

$$e = \phi - \phi_\chi \quad (39)$$

thus introducing a coupling between macro and micromorphic variables. Assuming isotropic material behavior for brevity, the additional contributions to the free energy can be taken as quadratic functions of e and $\nabla\phi_\chi$:

$$\psi(\underline{\tilde{\boldsymbol{\varepsilon}}}, T, q, \phi_\chi, \nabla\phi_\chi) = \psi^{(1)}(\underline{\tilde{\boldsymbol{\varepsilon}}}, T, q) + \psi^{(2)}(e = \phi - \phi_\chi, \nabla\phi_\chi, T), \quad \text{with} \quad (40)$$

$$\rho\psi^{(2)}(e, \nabla\phi_\chi, T) = \frac{1}{2}H_\chi(\phi - \phi_\chi)^2 + \frac{1}{2}A\nabla\phi_\chi \cdot \nabla\phi_\chi \quad (41)$$

where H_χ and A are the additional moduli introduced by the micromorphic model. After inserting the state laws (36)

$$\underline{\boldsymbol{a}} = \rho \frac{\partial\psi}{\partial\phi_\chi} = -H_\chi(\phi - \phi_\chi), \quad \underline{\boldsymbol{b}} = \rho \frac{\partial\psi}{\partial\nabla\phi_\chi} = A\nabla\phi_\chi \quad (42)$$

into the additional balance equation (28), the following partial differential equation is obtained, at least for a homogeneous material under isothermal conditions:

$$\phi = \phi_\chi - \frac{A}{H_\chi}\Delta\phi_\chi \quad (43)$$

where Δ is the Laplace operator. This type of equation is encountered at several places in the mechanics of generalized continua especially in the linear micromorphic theory (Mindlin 1964; Eringen 1999; Dillard et al. 2006) and in the so-called implicit gradient theory of plasticity and damage (Peerlings et al. 2001, 2004; Engelen et al. 2003). Note however that this equation corresponds to a special quadratic potential and represents the simplest micromorphic extension of the classical theory. It involves a characteristic length scale defined by:

$$l_c^2 = \frac{A}{H_\chi} \quad (44)$$

This length is real for positive values of the ratio A/H_χ . The additional material parameters H_χ and A are assumed to be positive in this work. This does not exclude a softening material behaviour that can be induced by the proper evolution of the internal variables (including $\phi = q$ itself).

Viscous Generalized Stress and Phase Field Model

Generalized stresses can also be associated with dissipation by introducing the viscous part a^v of a :

$$\underline{\tilde{\boldsymbol{\varepsilon}}} = \underline{\tilde{\boldsymbol{\varepsilon}}}^e + \underline{\tilde{\boldsymbol{\varepsilon}}}^p, \quad a = a^e + a^v \quad (45)$$

The entropy inequality (31) now becomes

$$\begin{aligned} & \left(\underline{\underline{\boldsymbol{\sigma}}} - \rho \frac{\partial \widehat{\psi}}{\partial \underline{\underline{\boldsymbol{\varepsilon}}^e}} \right) : \dot{\underline{\underline{\boldsymbol{\varepsilon}}^e} + \rho \left(\eta + \frac{\partial \widehat{\psi}}{\partial T} \right) \dot{T} + \left(a^e - \rho \frac{\partial \widehat{\psi}}{\partial \phi_\chi} \right) \dot{\phi}_\chi + \left(\underline{\underline{\mathbf{b}}} - \rho \frac{\partial \widehat{\psi}}{\partial \nabla \phi_\chi} \right) \cdot \nabla \dot{\phi}_\chi \\ & + \underline{\underline{\boldsymbol{\sigma}}} : \dot{\underline{\underline{\boldsymbol{\varepsilon}}^p} - \rho \frac{\partial \widehat{\psi}}{\partial q} \dot{q} + a^v \dot{\phi}_\chi - \frac{\underline{\underline{\mathbf{q}}}}{T} \cdot \nabla T \geq 0 \end{aligned} \quad (46)$$

Assuming that no dissipation is associated with the four first terms of the previous inequality, the following state laws are found

$$\underline{\underline{\boldsymbol{\sigma}}} = \rho \frac{\partial \widehat{\psi}}{\partial \underline{\underline{\boldsymbol{\varepsilon}}^e}, \quad \eta = -\frac{\partial \widehat{\psi}}{\partial T}, \quad R = -\rho \frac{\partial \widehat{\psi}}{\partial q} \quad (47)$$

$$a^e = \rho \frac{\partial \widehat{\psi}}{\partial \phi_\chi}, \quad \underline{\underline{\mathbf{b}}} = \rho \frac{\partial \widehat{\psi}}{\partial \nabla \phi_\chi} \quad (48)$$

and the residual dissipation is

$$D^{res} = \underline{\underline{\boldsymbol{\sigma}}} : \dot{\underline{\underline{\boldsymbol{\varepsilon}}^p} + R \dot{q} + a^v \dot{\phi}_\chi - \frac{\underline{\underline{\mathbf{q}}}}{T} \cdot \nabla T \geq 0 \quad (49)$$

Evolution rules for viscoplastic strain, internal variables, and the additional degrees of freedom can be derived from a dissipation potential $\Omega(\underline{\underline{\boldsymbol{\sigma}}}, R, a^v)$:

$$\dot{\underline{\underline{\boldsymbol{\varepsilon}}^p} = \frac{\partial \Omega}{\partial \underline{\underline{\boldsymbol{\sigma}}}, \quad \dot{q} = \frac{\partial \Omega}{\partial R}, \quad \dot{\phi}_\chi = \frac{\partial \Omega}{\partial a^v} \quad (50)$$

Convexity of the dissipation potential then ensures positivity of dissipation rate for any process.

Note that no dissipative part has been assigned to the generalized stress $\underline{\underline{\mathbf{b}}}$ since then exploitation of second principle does not seem to be straightforward. Instead, the total gradient $\nabla \phi_\chi$ can be split into elastic and plastic parts, as it will be done in section “[Elastic-Plastic Decomposition of Generalized Strains](#).”

Phase Field Model

The dissipation potential can be decomposed into the various contributions due to all thermodynamic forces. Let us assume for instance that the contribution of the viscous generalized stress a^v is quadratic:

$$\Omega(\underline{\underline{\boldsymbol{\sigma}}}, R, a^v) = \Omega_1(\underline{\underline{\boldsymbol{\sigma}}}, R) + \Omega_2(a^v), \quad \Omega_2(a^v) = \frac{1}{2\beta} a^{v2} \quad (51)$$

The use of the flow rule (50) and of the additional balance equation (28) then leads to

$$\beta \dot{\phi} = a^v = a - a^e = a - \rho \frac{\partial \widehat{\psi}}{\partial \phi_\chi} = \operatorname{div} \left(\rho \frac{\partial \widehat{\psi}}{\partial \nabla \phi_\chi} \right) - \rho \frac{\partial \widehat{\psi}}{\partial \phi_\chi} \quad (52)$$

One recognizes the Landau–Ginzburg equation that arises in phase field theories. The previous derivation of Landau–Ginzburg equation is due to Gurtin (1996), see also Ammar et al. (2009). The coupling with mechanics is straightforward according to this procedure and more general dissipative mechanisms can be put forward, see Rancourt et al. (2016).

Elastic–Plastic Decomposition of Generalized Strains

Instead of the previous decomposition of generalized stresses, the introduction of additional dissipative mechanisms can rely on the split of all strain measures into elastic and plastic parts:

$$\underline{\boldsymbol{\varepsilon}} = \underline{\boldsymbol{\varepsilon}}^e + \underline{\boldsymbol{\varepsilon}}^p, \phi_\chi = \phi_\chi^e + \phi_\chi^p, \underline{\boldsymbol{\kappa}} = \nabla \phi_\chi = \underline{\boldsymbol{\kappa}}^e + \underline{\boldsymbol{\kappa}}^p \quad (53)$$

The objectivity of ϕ_χ is required for such a unique decomposition to be defined. We do not require here that

$$\underline{\boldsymbol{\kappa}}^e = \nabla \phi_\chi^e, \underline{\boldsymbol{\kappa}}^p = \nabla \phi_\chi^p \quad (54)$$

although such a model also is possible, as illustrated by the *gradient of strain theory* put forward in Forest and Sievert (2003). The Clausius–Duhem inequality then writes

$$\begin{aligned} & \left(\underline{\boldsymbol{\sigma}} - \rho \frac{\partial \widehat{\psi}}{\partial \underline{\boldsymbol{\varepsilon}}^e} \right) : \dot{\underline{\boldsymbol{\varepsilon}}}^e + \rho \left(\eta + \frac{\partial \widehat{\psi}}{\partial T} \right) \dot{T} + \left(a - \rho \frac{\partial \widehat{\psi}}{\partial \phi_\chi} \right) \dot{\phi}_\chi^e + \left(\underline{\boldsymbol{b}} - \rho \frac{\partial \widehat{\psi}}{\partial \nabla \phi_\chi} \right) \cdot \dot{\underline{\boldsymbol{\kappa}}}^e \\ & + \underline{\boldsymbol{\sigma}} : \dot{\underline{\boldsymbol{\varepsilon}}}^p - \rho \frac{\partial \widehat{\psi}}{\partial q} \dot{q} + a \dot{\phi}_\chi^p + \underline{\boldsymbol{b}} \cdot \dot{\underline{\boldsymbol{\kappa}}}^p - \frac{q}{T} \cdot \nabla T \geq 0 \end{aligned} \quad (55)$$

Assuming that no dissipation is associated with the four first terms of the previous inequality, the following state laws are found

$$\underline{\boldsymbol{\sigma}} = \rho \frac{\partial \widehat{\psi}}{\partial \underline{\boldsymbol{\varepsilon}}^e}, \eta = -\frac{\partial \widehat{\psi}}{\partial T}, R = -\rho \frac{\partial \widehat{\psi}}{\partial q} \quad (56)$$

$$a = \rho \frac{\partial \widehat{\psi}}{\partial \phi_\chi}, \underline{\boldsymbol{b}} = \rho \frac{\partial \widehat{\psi}}{\partial \underline{\boldsymbol{\kappa}}^e} \quad (57)$$

and the residual dissipation is

$$D^{res} = \underline{\underline{\sigma}} : \underline{\underline{\dot{\epsilon}}}^p + R\dot{q} + a\dot{\phi}_\chi^p + \underline{\underline{b}} \cdot \underline{\underline{\dot{\kappa}}}^p - \frac{\underline{\underline{q}}}{T} \cdot \nabla T \geq 0 \quad (58)$$

Evolution rules for viscoplastic strain, internal variables, and the additional degrees of freedom can be derived from a dissipation potential $\Omega(\underline{\underline{\sigma}}, R, a, \underline{\underline{b}})$:

$$\underline{\underline{\dot{\epsilon}}}^p = \frac{\partial \Omega}{\partial \underline{\underline{\sigma}}}, \quad \dot{q} = \frac{\partial \Omega}{\partial R}, \quad \dot{\phi}_\chi^p = \frac{\partial \Omega}{\partial a}, \quad \underline{\underline{\dot{\kappa}}}^p = \frac{\partial \Omega}{\partial \underline{\underline{b}}} \quad (59)$$

As a result of the additional balance equation (28) combined with the previous state laws, the type of derived partial differential equation can be made more specific by adopting a quadratic free energy potential for $\underline{\underline{b}}$ (modulus A) and a quadratic dissipation potential with respect to a (parameter β). We obtain:

$$\beta \dot{\phi}_\chi = a + \beta \dot{\phi}_\chi^e = \text{div } A \underline{\underline{\kappa}} - \text{div } A \underline{\underline{\kappa}}^p + \beta \dot{\phi}_\chi^e \quad (60)$$

Decompositions of stresses and strains can also be mixed, for instance in the following way:

$$\underline{\underline{\epsilon}} = \underline{\underline{\epsilon}}^e + \underline{\underline{\epsilon}}^p, \quad a = a^e + a^v, \quad \underline{\underline{\kappa}} = \nabla \phi_\chi = \underline{\underline{\kappa}}^e + \underline{\underline{\kappa}}^p \quad (61)$$

based on which a constitutive theory can be built.

Application to a Continuum Damage Model for Single Crystals and Its Regularization

The micromorphic approach is now illustrated in the case of a constitutive model for damaging viscoplastic single crystals. The objective is to simulate crack initiation and propagation. The micromorphic model is used in order to obtain a regularized continuum damage formulation with a view to simulating mesh-independent crack propagation in single crystals.

Constitutive Equations

In the proposed crystal plasticity model taken from Marchal, et al. (2006a), viscoplasticity and damage are coupled by introducing an additional damage strain variable $\underline{\underline{\epsilon}}^d$, into the strain rate partition equation:

$$\underline{\underline{\dot{\epsilon}}} = \underline{\underline{\dot{\epsilon}}}^e + \underline{\underline{\dot{\epsilon}}}^p + \underline{\underline{\dot{\epsilon}}}^d \quad (62)$$

where $\tilde{\dot{\boldsymbol{\epsilon}}}^e$ and $\tilde{\dot{\boldsymbol{\epsilon}}}^p$ are the elastic and the plastic strain rates, respectively. The flow rule for plastic part is written at the slip system level by means of the orientation tensor $\tilde{\mathbf{m}}^s$:

$$\tilde{\mathbf{m}}^s = \frac{1}{2} (\underline{\mathbf{n}}^s \otimes \underline{\mathbf{l}}^s + \underline{\mathbf{l}}^s \otimes \underline{\mathbf{n}}^s) \quad (63)$$

where $\underline{\mathbf{n}}^s$ is the normal to the plane of slip system s and $\underline{\mathbf{l}}^s$ stands for the corresponding slip direction. Then, plastic strain rate reads:

$$\tilde{\dot{\boldsymbol{\epsilon}}}^p = \sum_{s=1}^{N_{slip}} \dot{\gamma}^s \tilde{\mathbf{m}}^s \quad (64)$$

where N_{slip} is the total number of slip systems. The flow rule on slip system s is a classical Norton rule with threshold:

$$\dot{\gamma}^s = \left\langle \frac{|\tau^s - x^s| - r^s}{K} \right\rangle^n \text{sign}(\tau^s - x^s) \quad (65)$$

where r^s and x^s are the variables for isotropic and kinematic hardening respectively and K and n are viscosity material parameters to be identified from experimental curves.

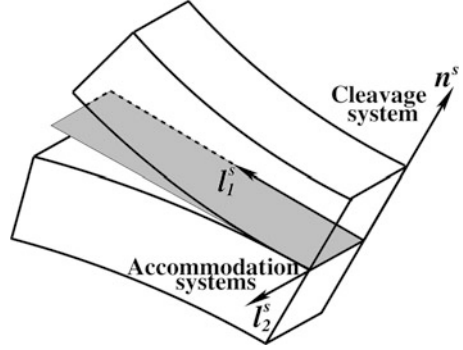
Material separation is assumed to take place w.r.t. specific crystallographic planes, like cleavage planes in single crystals. The word *cleavage* is written in a more general sense that its original meaning in physical metallurgy associated with brittle fracture of non-f.c.c. crystals. In the continuum mechanical model, *cleavage* means cracking along a specific crystallographic plane as it is often observed in low cycle fatigue of f.c.c. crystals like single crystal nickel–base superalloys. The damage strain $\tilde{\dot{\boldsymbol{\epsilon}}}^d$ is decomposed in the following crystallographic contributions:

$$\tilde{\dot{\boldsymbol{\epsilon}}}^d = \sum_{s=1}^{N_{damage}} \dot{\delta}_c^s \underline{\mathbf{n}}_d^s \otimes \underline{\mathbf{n}}_d^s + \dot{\delta}_1^s \underline{\mathbf{n}}_d^s \otimes \underline{\mathbf{l}}_{d_1}^s + \dot{\delta}_2^s \underline{\mathbf{n}}_d^s \otimes \underline{\mathbf{l}}_{d_2}^s \quad (66)$$

where $\dot{\delta}^s$, $\dot{\delta}_1^s$ and $\dot{\delta}_2^s$ are the strain rates for mode I, mode II and mode III crack growth, respectively and N_{damage}^d stands for the number of damage planes which are fixed crystallographic planes depending on the crystal structure. Cleavage damage is represented by the opening $\dot{\delta}_c^s$ of crystallographic cleavage planes with the normal vector $\underline{\mathbf{n}}^s$. Additional damage systems must be introduced for the in-plane accommodation along orthogonal directions $\underline{\mathbf{l}}_1^s$ and $\underline{\mathbf{l}}_2^s$, once cleavage has started (Fig. 4). Three damage criteria are associated to one cleavage and two accommodation systems:

$$f_c^s = \left| \underline{\mathbf{n}}_d^s \cdot \underline{\boldsymbol{\sigma}} \cdot \underline{\mathbf{n}}_d^s \right| - Y_c^s \quad (67)$$

Fig. 4 Illustration of the cleavage and two accommodation systems to be associated to the crystallographic planes



$$f_i^s = \left| \underline{n}_d^s \cdot \underline{\sigma} \cdot \underline{l}_{d_i}^s \right| - Y_i^s \quad (i = 1, 2) \quad (68)$$

The critical normal stress Y^s for damage decreases as δ increases:

$$Y_c^s = Y_0^s + H \delta_c^s, \quad Y_i^s = Y_0^s + H \delta_i^s \quad (69)$$

where Y_0^s is the initial damage stress (usually coupled to plasticity) and H is a negative modulus which controls material softening due to damage. Finally, evolution of damage is given by the following equations;

$$\dot{\delta}_c^s = \left\langle \frac{f_c^s}{K_d} \right\rangle^{n_d} \text{sign} \left(\underline{n}_d^s \cdot \underline{\sigma} \cdot \underline{n}_d^s \right) \quad (70)$$

$$\dot{\delta}_i^s = \left\langle \frac{f_i^s}{K_d} \right\rangle^{n_d} \text{sign} \left(\underline{n}_d^s \cdot \underline{\sigma} \cdot \underline{l}_{d_i}^s \right) \quad (71)$$

where K_d and n_d are material parameters.

These equations hold for all conditions except when the crack is closed ($\delta_c^s < 0$) and compressive forces are applied ($\underline{n}_d^s \cdot \underline{\sigma} \cdot \underline{n}_d^s < 0$). In this case, damage evolution stops ($\dot{\delta}_c^s = \dot{\delta}_i^s = 0$), corresponding to the unilateral damage conditions.

Note that the damage variables δ introduced in the model differ from the usual corresponding variables of standard continuum damage mechanics that vary from 0 to 1. In contrast, the δ s are strain-like quantities that can ever increase.

Coupling between plasticity and damage is generated through initial damage stress Y_0 in (69) which is controlled by cumulative slip variable γ_{cum} :

$$\dot{\gamma}_{cum} = \sum_{s=1}^{N_{slips}} |\dot{\gamma}^s| \quad (72)$$

Then, Y_0 takes the form:

$$Y_0^s = \sigma_n^c e^{-d\gamma_{cum}} + \sigma_{ult} \quad (73)$$

This formulation suggests an exponential decaying regime from a preferably high initial cleavage stress value σ_n^c , to an ultimate stress, σ_{ult} which is close but not equal to zero for numerical reasons and d is a material constant.

This model, complemented by the suitable constitutive equations for viscoplastic strain, has been used for the simulation of crack growth under complex cyclic loading at high temperature (Marchal et al. 2006a). Significant mesh dependency of results was found Marchal et al. (2006b).

In the present work, the model is further developed by switching from classical to microdamage continuum in order to assess the regularization capabilities of a higher order theory. The coupling of the model with microdamage theory is achieved by introducing a cumulative damage variable calculated from the damage systems and a new threshold function $Y_0(\delta, \gamma_{cum})$:

$$\dot{\delta}_{cum} = \sum_{s=1}^{N_{planes}} \dot{\delta}^s, \quad \text{where } \dot{\delta}^s = \left| \dot{\delta}_c^s \right| + \left| \dot{\delta}_1^s \right| + \left| \dot{\delta}_2^s \right| \quad (74)$$

$$Y_0 = \sigma_n^c e^{-d\gamma_{cum} - H\delta_{cum}} + \sigma_{ult} \quad (75)$$

where the modulus H accounts for damage induced softening and σ_{ult} is a ultimate stress.

Microdamage Continuum

In microdamage theory, the introduced microvariable is a scalar microdamage field δ_χ :

$$DOF = \{ \underline{\mathbf{u}}, \delta_\chi \} \quad STRAIN = \{ \underline{\boldsymbol{\varepsilon}}, \delta_\chi, \nabla \delta_\chi \} \quad (76)$$

The power of internal forces is extended as

$$p^{(i)} = \underline{\boldsymbol{\sigma}} : \underline{\dot{\boldsymbol{\varepsilon}}} + a \dot{\delta}_\chi + \underline{\mathbf{b}} \cdot \nabla \dot{\delta}_\chi \quad (77)$$

where generalized stresses $a, \underline{\mathbf{b}}$ have been introduced. The generalized balance equations are:

$$\text{div } \underline{\boldsymbol{\sigma}} = 0, \quad a = \text{div } \underline{\mathbf{b}} \quad (78)$$

The free energy density is taken as a quadratic potential in the elastic strain, damage δ , relative damage $\delta - \delta_\chi$ and microdamage gradient $\nabla\delta_\chi$:

$$\rho\psi = \frac{1}{2}\underline{\underline{\boldsymbol{\varepsilon}}^e : \underline{\underline{\boldsymbol{c}}} : \underline{\underline{\boldsymbol{\varepsilon}}^e} + \frac{1}{2} \sum_{s=1}^{N_{\text{damage}}} H\delta_s^2 + \frac{1}{2}H_\chi(\delta - \delta_\chi)^2 + \frac{1}{2}A\nabla\delta_\chi \cdot \nabla\delta_\chi \quad (79)$$

where H , H_χ and A are scalar material constants. The tensor of elastic moduli is called $\underline{\underline{\boldsymbol{c}}}$. Then, the elastic response of the material becomes:

$$\underline{\underline{\boldsymbol{\sigma}}} = \rho \frac{\partial\psi}{\partial\underline{\underline{\boldsymbol{\varepsilon}}^e} = \underline{\underline{\boldsymbol{c}}} : \underline{\underline{\boldsymbol{\varepsilon}}^e \quad (80)$$

The generalized stresses read:

$$\underline{\underline{\boldsymbol{a}}} = \rho \frac{\partial\psi}{\partial\delta} = -H_\chi(\delta - \delta_\chi), \quad \underline{\underline{\boldsymbol{b}}} = A\nabla\delta_\chi \quad (81)$$

and the driving force for damage can be derived as:

$$Y^s = \rho \frac{\partial\psi}{\partial\delta^s} = H\delta^s + H_\chi(\delta^s - \delta_\chi) \quad (82)$$

The damage criterion now is:

$$f^s = \left| \underline{\underline{\boldsymbol{n}}}^s \cdot \underline{\underline{\boldsymbol{\sigma}}} \cdot \underline{\underline{\boldsymbol{n}}}^s \right| - Y_0 - Y^s = 0 \quad (83)$$

Substituting the linear constitutive equations for generalized stresses into the additional balance equation (78), assuming homogeneous material properties, leads to the following partial differential equation for the microdamage

$$\delta_\chi - \frac{A}{H_\chi} \Delta\delta_\chi = \delta \quad (84)$$

where the macrodamage δ acts as a source term. Exactly this type of Helmholtz equation has been postulated in the so-called implicit gradient theory of plasticity and damage (Peerlings et al. 2001, 2004; Engelen et al. 2003; Germain et al. 2007), where the microvariables are called non local variables and where the generalized stresses $\underline{\underline{\boldsymbol{a}}}$ and $\underline{\underline{\boldsymbol{b}}}$ are not explicitly introduced (see Forest (2009) and Dillard et al. (2006) for the analogy between this latter approach and the micromorphic theory).

The solution of the problem of failure of a 1D bar in tension/compression was treated by Aslan and Forest (2009). The characteristic size of the damage zone was shown to be controlled by the characteristic length corresponding to the inverse of

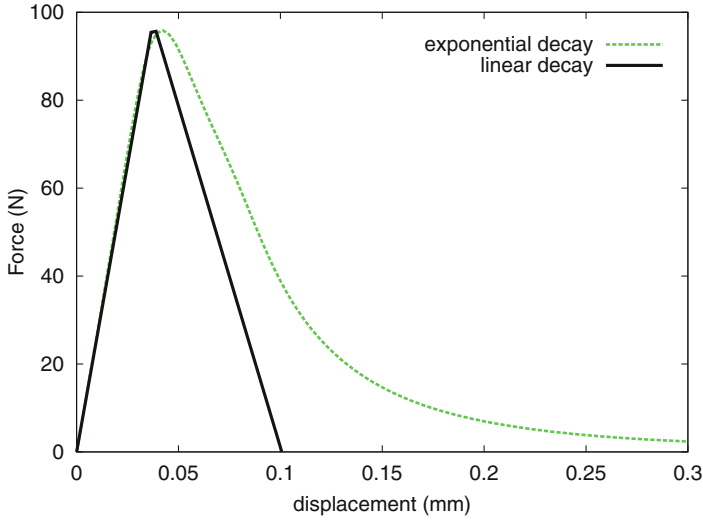


Fig. 5 Comparison between force vs. displacement diagram of a 1D softening rod for linear and exponential decay

$$\omega = \sqrt{|H H_\chi / A (H + H_\chi)|} \quad (85)$$

In comparison with the standard strain gradient approaches (Peerlings et al. 2001; Germain et al. 2007), microdamage theory eliminates the final fracture problem without any modification to the damage function, since there exists no direct coupling between the force stress $\underline{\sigma}$ and the generalized stresses, \underline{a} and \underline{b} . For a better representation of a cracked element, an exponential drop is used for both the damage threshold Y_0 and the modulus A , since the element should become unable to store energy due to the generalized stresses when broken (see Fig. 5):

$$Y_0 = \sigma_n^c e^{-H\delta} + \sigma_{ult}, \quad \underline{b} = A e^{-H\delta} \nabla \chi \delta \quad (86)$$

Numerical Examples

As a 2D example, a single crystal CT-like specimen under tension is analysed. The corresponding FE mesh is given in Fig. 6. Analyses are performed for two different crack widths, obtained by furnishing different material parameters which control the characteristic length (Fig. 7). The propagation of a crack, corresponding

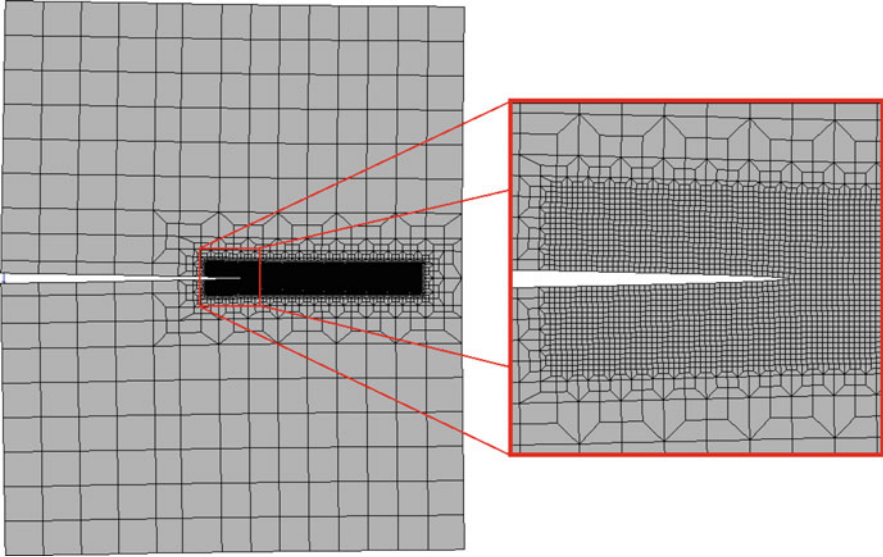


Fig. 6 FE mesh of a CT-like specimen

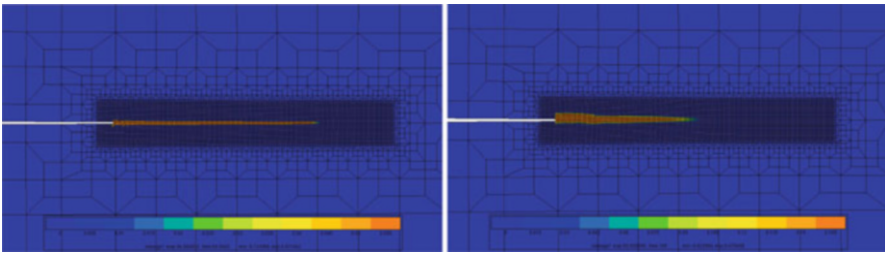


Fig. 7 Crack growth in a 2D single crystal CT-like specimen with a single cleavage plane aligned through the horizontal axis under vertical tension. Field variable δ (Left) $A = 100 \text{ MPa}\cdot\text{mm}^2$, $H = -20000 \text{ MPa}$, $H_\chi = 30,000 \text{ MPa}$, (Right) $A = 150 \text{ MPa}\cdot\text{mm}^2$, $H = -10000 \text{ MPa}$, $H_\chi = 30000 \text{ MPa}$

stress fields and the comparison with classical elastic solutions are given in Fig. 8. This comparison shows that the microdamage model is able to reproduce the stress concentration at the crack tip except very close to the crack tip where finite stress values are predicted.

The Micromorphic Approach at Finite Deformations

Extensions of the previous systematic micromorphic approach are presented for finite deformations along the lines of Forest (2016). They generalize some aspects

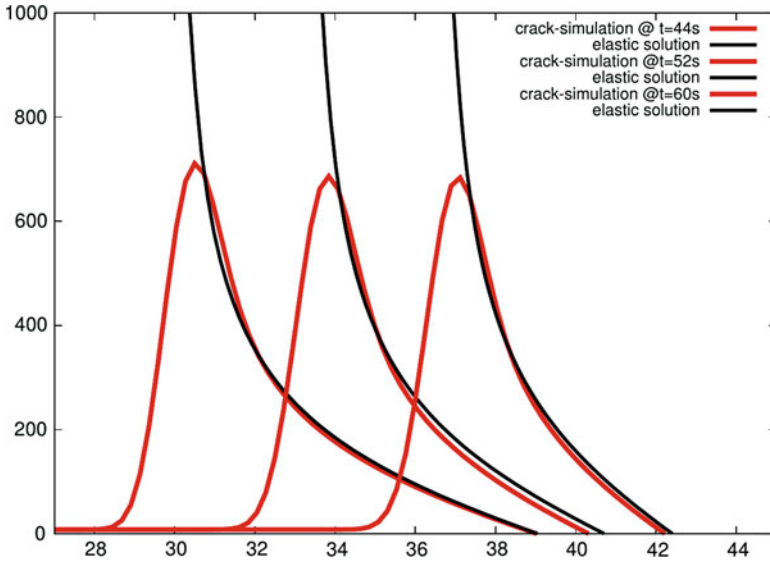


Fig. 8 Mode I stress profile (in MPa) vs. position along the ligament (in mm) at three distinct crack propagation steps: micromorphic solution compared to the linear elastic one

of the finite deformation micromorphic theory presented in section “[The Micromorphic Theory After Eringen and Mindlin and Its Extension to Plasticity](#)”.

Micromorphic and Gradient Hyperelasticity

Nonlinearity arises not only from nonlinear material response but also from the consideration of finite deformations. The impact of finite strains on regularisation operators is largely unexplored. It is first illustrated in the pure hyperelastic case, leaving aside for a moment the inclusion of plastic effects. It is recalled that the Lagrangian coordinates of the material points are denoted by \underline{X} on the reference configuration Ω_0 , whereas their positions in the current configuration are called \underline{x} . The gradient operators with respect to Lagrangian (reference) and Eulerian (current) coordinates are denoted by $\nabla_{\underline{X}}$ and $\nabla_{\underline{x}}$, respectively. The deformation gradient is $\underline{F} = \underline{\mathbf{1}} + \underline{u} \otimes \nabla_{\underline{X}}$ where the displacement vector is the function $\underline{u}(\underline{X}, t)$. The initial and current mass density functions are $\rho_0(\underline{X}, t)$ and $p(\underline{x}, t)$, respectively.

It is appropriate here to recall the Eulerian balance equation and boundary conditions in the case of a scalar micromorphic degree of freedom, in the absence of body forces:

$$\operatorname{div} \underline{\underline{\sigma}} = 0, a = \operatorname{div} \underline{b}, \forall \underline{x} \in \Omega, \underline{t} = \underline{\underline{\sigma}} \cdot \underline{n}, a_c = \underline{b} \cdot \underline{n}, \forall \underline{x} \in \partial\Omega \quad (87)$$

Finite Microstrain Tensor Model

The microstrain model was introduced in Table 1 as a reduced micromorphic model where Eringen's microdeformation tensor is taken as symmetric, as proposed in Forest and Sievert (2006). The additional degrees of freedom are the six components of a microstrain tensor $\underline{\underline{\chi}}$, a second order symmetric tensor associated with the right Cauchy-Green strain measure of the micromorphic deformation. Accordingly, the microstrain is regarded here as a Lagrangian variable. The Lagrangian power density of internal forces takes the form:

$$p_0^{(i)} = \frac{1}{2} \underline{\underline{\Pi}} : \underline{\underline{\dot{C}}} + \underline{\underline{a}}_0 : \underline{\underline{\dot{\chi}}} + \underline{\underline{b}}_0 \cdot (\underline{\underline{\dot{\chi}}} \otimes \nabla_X) = J p^{(i)} \quad (88)$$

$p^{(i)}$ being the Eulerian internal power density and the Jacobian $J = \det \underline{\underline{F}}$. The right Cauchy-Green tensor is $\underline{\underline{C}} = \underline{\underline{F}}^T \cdot \underline{\underline{F}}$ and $\underline{\underline{\Pi}}$ is the Piola stress tensor. The method of virtual power can be used to show that the generalized Lagrangian stress tensors, a second and a third rank tensor, fulfill the following balance equation:

$$\text{Div } \underline{\underline{b}}_0 = \underline{\underline{a}}_0 \quad (89)$$

in addition to the balance of momentum equation

$$\text{Div } \underline{\underline{F}} \cdot \underline{\underline{\Pi}} = 0 \quad (90)$$

in the absence of body and inertial forces, the divergence operator Div being computed with respect to Lagrangian coordinates. The corresponding Eulerian forms of the balance equations are

$$\text{div } \underline{\underline{\sigma}} = 0, \quad \text{div } \underline{\underline{b}} = \underline{\underline{a}} \quad (91)$$

with the Neumann boundary conditions for tractions and double tractions: $\underline{\underline{t}} = \underline{\underline{\sigma}} \cdot \underline{\underline{n}}, \underline{\underline{b}} = \underline{\underline{b}} \cdot \underline{\underline{n}}$, involving the normal vector of the current surface element. The relations between the Lagrangian and Eulerian generalized stress tensors

$$\underline{\underline{a}}_0 = J \underline{\underline{a}}, \quad \underline{\underline{b}}_0 = J \underline{\underline{b}} \cdot \underline{\underline{F}}^{-T} = J \underline{\underline{F}}^{-1} \cdot \underline{\underline{b}} \quad (92)$$

The hyperelastic free energy density function is $\psi_0 \left(\underline{\underline{C}}, \underline{\underline{\chi}}, \underline{\underline{\chi}} \otimes \nabla_X \right)$ and the stress-strain relations read:

$$\underline{\underline{\Pi}} = 2\rho_0 \frac{\partial \psi_0}{\partial \underline{\underline{C}}}, \quad \underline{\underline{a}}_0 = \rho_0 \frac{\partial \psi_0}{\partial \underline{\underline{\chi}}}, \quad \underline{\underline{b}}_0 = \rho_0 \frac{\partial \psi_0}{\partial \underline{\underline{\chi}} \otimes \nabla_X} \quad (93)$$

The relative strain tensor $\underline{\underline{e}} = \underline{\underline{C}} - \underline{\underline{\chi}}$ was defined in Forest and Sievert (2006) and measures the difference between macro and micro-strain. As an example, the following potential is proposed:

$$\rho_0 \psi_0 = \rho_0 \psi_{ref}(\underline{\underline{C}}) + \frac{1}{2} H_\chi (\underline{\underline{C}} - \underline{\underline{\chi}})^2 + \frac{1}{2} \underline{\underline{\chi}} \otimes \nabla_X : \underline{\underline{A}} : \underline{\underline{\chi}} \otimes \nabla_X \quad (94)$$

where a penalty modulus H_χ is introduced and where ψ_{ref} is a standard hyperelastic strain energy potential (neo-Hookean, etc.). The higher order term involves a sixth-rank tensor of elasticity moduli which is symmetric and assumed definite positive, see Auffray et al. (2013). The stress-strain relations (93) become:

$$\underline{\underline{\Pi}} = 2\rho_0 \frac{\partial \psi_{ref}}{\partial \underline{\underline{C}}} + H_\chi (\underline{\underline{C}} - \underline{\underline{\chi}}), \underline{\underline{a}}_0 = -H_\chi (\underline{\underline{C}} - \underline{\underline{\chi}}), \underline{\underline{b}}_0 = \underline{\underline{A}} : (\underline{\underline{\chi}} \otimes \nabla_X) \quad (95)$$

Note that the classical hyperelastic relation is complemented by a coupling term involving the microstrain tensor. Taking the balance equation (90) into account, the set of p.d.e. for the microstrain components is found to be:

$$\underline{\underline{C}} = \underline{\underline{\chi}} - \frac{1}{H_\chi} \text{Div} \left(\underline{\underline{A}} : (\underline{\underline{\chi}} \otimes \nabla_X) \right) \quad (96)$$

The associated regularisation operator is now given in the simplified case where the sixth rank tensor of higher order moduli is assumed to be the identity multiplied by the single modulus A :

$$\text{Op} = 1 - \frac{A}{H_\chi} \Delta_X \quad (97)$$

It involves the Lagrangian Laplace operator $\Delta_X(\bullet) = (\bullet)_{,KK}$ in a Cartesian frame where capital indices refer to Lagrange coordinates and the comma to partial derivation. It is linear w.r.t. Lagrangian coordinates but the full problem is of course highly non linear. The associated Eulerian partial differential operator is nonlinear in the form: $\chi_{IJ, KK} = \chi_{IJ, KI} F_{kK} F_{IK}$, where small index letters refer to the current coordinate system.

An Eulerian formulation of the proposed constitutive equations is possible. It will be illustrated in the next section in the case of a scalar microstrain variable for the sake of brevity.

Equivalent Microstrain Model at Finite Deformation

According to the micromorphic approach, it is not necessary to consider the full microstrain tensor. Instead, the set of degrees of freedom of a reduced micromorphic

model can be the usual displacement vector and a scalar microstrain variable $\chi(\underline{\mathbf{X}}, t)$. The latter variable is assumed to be a Lagrangian quantity, invariant w.r.t. change of observer. The free energy density is a function of the following argument $\psi_0(\underline{\mathbf{C}}, \chi, \nabla_x \chi)$. The corresponding hyperelastic state laws fulfilling the Clausius–Duhem inequality take the form:

$$\underline{\underline{\boldsymbol{\Pi}}} = 2\rho_0 \frac{\partial \psi_0}{\partial \underline{\underline{\mathbf{C}}}}, \quad a_0 = \rho_0 \frac{\partial \psi_0}{\partial \chi}, \quad \underline{\underline{\mathbf{b}}}_0 = \rho_0 \frac{\partial \psi_0}{\partial \nabla_x \chi} \quad (98)$$

As an example, the following Lagrangian potential is proposed:

$$\rho_0 \psi_0 = \rho_0 \psi_{ref}(\underline{\underline{\mathbf{C}}}) + \frac{1}{2} H_\chi (C_{eq} - \chi)^2 + \frac{1}{2} \nabla_x \chi \cdot \underline{\underline{\mathbf{A}}} \cdot \nabla_x \chi \quad (99)$$

The microstrain is compared to some equivalent strain measure, C_{eq} , function of the invariants of $\underline{\underline{\mathbf{C}}}$. The stress–strain relations (98) become:

$$\underline{\underline{\boldsymbol{\Pi}}} = 2\rho_0 \frac{\partial \psi_{ref}}{\partial \underline{\underline{\mathbf{C}}}} + H_\chi (C_{eq} - \chi) \frac{\partial C_{eq}}{\partial \underline{\underline{\mathbf{C}}}}, \quad a_0 = -H_\chi (C_{eq} - \chi), \quad \underline{\underline{\mathbf{b}}}_0 = \underline{\underline{\mathbf{A}}} \cdot \nabla_x \chi \quad (100)$$

Taking the balance equation $a_0 = \nabla_x \cdot \underline{\underline{\mathbf{b}}}_0$ into account, the p.d.e. governing χ is:

$$C_{eq} = \chi - \frac{1}{H_\chi} \text{Div}(\underline{\underline{\mathbf{A}}} \cdot \nabla_x \chi) \quad (101)$$

Let us mention the corresponding isotropic form, when $\underline{\underline{\mathbf{A}}} = A \mathbf{1}$:

$$C_{eq} = \chi - \frac{A}{H_\chi} \Delta_x \chi \quad (102)$$

where Δ_x is the Laplace operator with respect to Lagrangian coordinates. An example of equivalent strain measure which the microstrain is compared with is

$$C_{eq} = \sqrt{\underline{\underline{\mathbf{C}}} : \underline{\underline{\mathbf{C}}}}, \quad \frac{\partial C_{eq}}{\partial \underline{\underline{\mathbf{C}}}} = \underline{\underline{\mathbf{C}}} / C_{eq} \quad (103)$$

The penalty modulus H_χ constrains the microstrain degree of freedom to remain close to the equivalent strain measure C_{eq} . In the limit of large values of H_χ , the generalized stress a_0 can be regarded as a Lagrange multiplier enforcing the constraint. As a result, the microstrain gradient becomes equal to the equivalent strain gradient and the micromorphic model degenerates into a gradient model,

see Forest and Sab (2017) for a detailed description of internal constraints in micromorphic theories.

The formulation of a constitutive law based on Eulerian strain measures, $\underline{\underline{\mathbf{B}}} = \underline{\underline{\mathbf{F}}} \cdot \underline{\underline{\mathbf{F}}}^T$ and $\nabla_x \chi$, is now envisaged. It relies on the choice of a free energy potential $\rho\psi(\underline{\underline{\mathbf{B}}}, \chi, \nabla_x \chi)$. Galilean invariance of the constitutive law requires that this function fulfills the following conditions:

$$\psi\left(\underline{\underline{\mathbf{Q}}} \cdot \underline{\underline{\mathbf{B}}} \cdot \underline{\underline{\mathbf{Q}}}^T, \chi, \underline{\underline{\mathbf{Q}}} \cdot \nabla_x \chi\right) = \psi\left(\underline{\underline{\mathbf{B}}}, \chi, \nabla_x \chi\right) \quad (104)$$

for all constant orthogonal transformations $\underline{\underline{\mathbf{Q}}}$. This amounts to stating isotropy of the function ψ . Representation theorems are available for such functions, $\psi(\underline{\underline{\mathbf{B}}}, \chi, \nabla_x \chi) = \psi(B_1, B_2, B_3, \chi, \|\nabla_x \chi\|)$, where the B_i are the eigenvalues of $\underline{\underline{\mathbf{B}}}$. The Cauchy stress tensor, $\underline{\underline{\boldsymbol{\sigma}}}$, is known then to commute with $\underline{\underline{\mathbf{B}}}$ such that:

$$\underline{\underline{\boldsymbol{\sigma}}} : \underline{\underline{\mathbf{D}}} = \left(\underline{\underline{\boldsymbol{\sigma}}} \cdot \underline{\underline{\mathbf{B}}}^{-1}\right) : \left(\underline{\underline{\dot{\mathbf{F}}}} \cdot \underline{\underline{\mathbf{F}}}^T\right) = \left(\underline{\underline{\mathbf{B}}}^{-1} \cdot \underline{\underline{\boldsymbol{\sigma}}}\right) : \left(\underline{\underline{\mathbf{F}}} \cdot \underline{\underline{\dot{\mathbf{F}}}}^T\right) = \frac{1}{2} \left(\underline{\underline{\mathbf{B}}}^{-1} \cdot \underline{\underline{\boldsymbol{\sigma}}}\right) : \underline{\underline{\dot{\mathbf{B}}}} \quad (105)$$

where the strain rate tensor, $\underline{\underline{\mathbf{D}}}$, is the symmetric part of the velocity gradient, $\underline{\underline{\dot{\mathbf{F}}}} \cdot \underline{\underline{\mathbf{F}}}^{-1}$. The hyperelastic state laws then take the form:

$$\underline{\underline{\boldsymbol{\sigma}}} = 2\underline{\underline{\mathbf{B}}} \cdot \rho \frac{\partial \psi}{\partial \underline{\underline{\mathbf{B}}}}, \quad a = \rho \frac{\partial \psi}{\partial \chi}, \quad \underline{\underline{\mathbf{b}}} = \rho \frac{\partial \psi}{\partial \nabla_x \chi} \quad (106)$$

As an example, the following Eulerian potential is proposed:

$$\rho\psi = \rho\psi_{ref}(\underline{\underline{\mathbf{B}}}) + \frac{1}{2}H_\chi(B_{eq} - \chi)^2 + \frac{1}{2}A\nabla_x \chi : \nabla_x \chi, \quad (107)$$

where ψ_{ref} refers to a standard isotropic elasticity potential from the classical finite elasticity theory. Note that $B_{eq} = C_{eq}$ since $\underline{\underline{\mathbf{B}}}$ and $\underline{\underline{\mathbf{C}}}$ share the same eigenvalues. The state laws (106) become:

$$\underline{\underline{\boldsymbol{\sigma}}} = 2\rho \frac{\partial \psi_{ref}}{\partial \underline{\underline{\mathbf{B}}}} + H_\chi (B_{eq} - \chi) \frac{\partial B_{eq}}{\partial \underline{\underline{\mathbf{B}}}}, \quad a = -H_\chi (B_{eq} - \chi), \quad \underline{\underline{\mathbf{b}}} = A\nabla_x \chi \quad (108)$$

The Eulerian regularisation operator follows from (91):

$$B_{eq} = \chi - \frac{A}{H_\chi} \Delta_x \chi \quad (109)$$

where Δ_x is the Laplace operator with respect to the Eulerian coordinates.

It is essential to notice that the isotropic regularisation operators (102) and (109) are distinct. For, if the Lagrangian higher order elastic law is linear with respect to the constitutive quantities involved, the deduced Eulerian law is NOT linear:

$$\underline{b}_0 = A \nabla_x \chi \quad \Rightarrow \quad \underline{b} = J^{-1} \underline{F} \cdot \underline{b}_0 = A J^{-1} \underline{B} \cdot \nabla_x \chi \quad (110)$$

so that the Eulerian regularisation operator will not be linear, i.e. different from (109).

Finite Deformation Micromorphic Elastoviscoplasticity Using an Additive Decomposition of a Lagrangian Strain

The most straightforward extension of the previous framework to viscoplasticity is to introduce a finite plastic strain measure in the decomposition of a Lagrangian total strain tensor. Such Lagrangian formulations of elastoviscoplasticity involve the additive decomposition of some Lagrangian total strain measure into elastic and viscoplastic parts:

$$\underline{E}_h = h(\underline{C}) = \underline{E}_h^e + \underline{E}_h^p \quad (111)$$

Many choices are possible for the invertible function h with restrictions ensuring that \underline{E}_h is a strain measure (symmetric, vanishing for rigid body motions, differentiable at 0 so that the tangent is the usual small strain tensor $\underline{\epsilon}$, used before, see Besson et al. 2009). Seth-Hill's strain measures are obtained for power law functions such that: $\underline{E}_m = \frac{1}{m} (\underline{C}^{\frac{m}{2}} - \underline{1})$, for $m > 0$, $\underline{E}_0 = \log \underline{C}^{\frac{1}{2}}$, the latter being the Lagrangian logarithmic strain. The case $m = 2$ corresponds to the Green–Lagrange strain measure for which this finite deformation theory was first formulated by Green and Naghdi (see Lee and Germain 1972; Bertram 2012) for the pros and the cons of various such formulations. This Lagrangian formulation is preferable to Eulerian ones based on corresponding Eulerian strain measures in order not to limit the approach to isotropic material behaviour (Simo and Miehe 1992). The additive decomposition of the Lagrangian logarithmic strain is put forward in the computational plasticity strategies developed in Miehe et al. (2002) and Helfer (2015). However, there is generally no physical motivation for the selection of one or another Lagrangian strain measure within this framework. In addition, this approach favours one particular reference configuration for which the corresponding strain is decomposed into elastic and plastic parts, again without clear physical argument. Changes of reference configuration lead to complex hardly interpretable

transformation rules for the plastic strain variables, see Shotov and Ihlemann (2014) for a comparison of finite deformation constitutive laws with respect to this issue. Note also that limitations arise from using a symmetric plastic strain variable \mathbf{E}_h^p especially when plastic spin relations are needed for anisotropic materials (Bertram 2012). This framework was applied to micromorphic and gradient plasticity and damage theories in Geers (2004), Horak and Jirasek (2013), and Miehe (2014).

A Lagrangian conjugate stress tensor $\underline{\underline{\Pi}}_{\sim h}$ is defined for each strain measure \mathbf{E}_h such that

$$\frac{1}{2} \underline{\underline{\Pi}}_{\sim} : \dot{\mathbf{C}}_{\sim} = \underline{\underline{\Pi}}_{\sim h} : \dot{\mathbf{E}}_{\sim h} \text{ with } \underline{\underline{\Pi}}_{\sim h} = \underline{\underline{\Pi}}_{\sim} : \left(\frac{\partial h}{\partial \mathbf{C}_{\sim}} \right)^{-1} \quad (112)$$

The power density of internal forces is:

$$p_0^{(i)} = \underline{\underline{\Pi}}_{\sim h} : \dot{\mathbf{E}}_{\sim h} + a_0 \dot{\chi} + \mathbf{b}_0 \cdot \nabla_X \dot{\chi}$$

and the free energy density function has the following arguments: $\psi_0(\mathbf{E}_h^e, q, \chi, \mathbf{K} := \nabla X \chi)$. The dissipation inequality then reads:

$$\begin{aligned} & \left(\underline{\underline{\Pi}}_{\sim h} - \rho_0 \frac{\partial \psi_0}{\partial \mathbf{E}_h^e} \right) : \dot{\mathbf{E}}_{\sim h} + \left(a - \rho_0 \frac{\partial \psi_0}{\partial \chi} \right) \dot{\chi} + \left(\mathbf{b}_0 - \rho_0 \frac{\partial \psi_0}{\partial \nabla_X \chi} \right) \cdot \dot{\mathbf{K}} \\ & + \underline{\underline{\Pi}}_{\sim h} : \dot{\mathbf{E}}_{\sim h}^p - \rho_0 \frac{\partial \psi_0}{\partial q} \dot{q} \geq 0 \end{aligned} \quad (113)$$

from which the following state laws are selected:

$$\underline{\underline{\Pi}}_{\sim h} = \rho_0 \frac{\partial \psi_0}{\partial \mathbf{E}_h^e}, \quad a = \rho_0 \frac{\partial \psi_0}{\partial \text{ffl}}, \quad \mathbf{b}_0 = \rho_0 \frac{\partial \psi_0}{\partial \nabla_X \chi}, \quad R = \rho_0 \frac{\partial \psi_0}{\partial q} \quad (114)$$

The flow and hardening rules can be determined from the suitable choice of dissipation potential $\Omega(\underline{\underline{\Pi}}_{\sim h}, R)$:

$$\dot{\mathbf{E}}_{\sim h}^p = \frac{\partial \Omega}{\partial \underline{\underline{\Pi}}_{\sim h}}, \quad \dot{q} = -\frac{\partial \Omega}{\partial R} \quad (115)$$

The existence of such a dissipation potential is not necessary but assumed in the whole chapter for convenience. Alternative methods of exploitation of the

dissipation principle exist for micromorphic continua, for instance based on the extended Liu procedure (Ván et al. 2014; Berezovski et al. 2014).

Two straightforward extensions of the micromorphic approach to finite strain viscoplasticity based on an additive decomposition of a Lagrangian strain measure are presented:

$$\rho_0 \psi_0 = \frac{1}{2} \underline{\underline{\mathbf{E}}}_h^e : \underline{\underline{\mathbf{c}}} : \underline{\underline{\mathbf{E}}}_h^e + \rho_0 \psi_q(q) + H_\chi (E_{heq} - \chi)^2 + \rho_0 \psi_\nabla(\underline{\underline{\mathbf{K}}}) \quad (116)$$

where E_{heq} is an equivalent total strain measure, or, alternatively,

$$\rho_0 \psi_0 = \frac{1}{2} \underline{\underline{\mathbf{E}}}_h^e : \underline{\underline{\mathbf{c}}} : \underline{\underline{\mathbf{E}}}_h^e + \rho_0 \psi_q(q) + H_\chi (p - \chi)^2 + \rho_0 \psi_\nabla(\underline{\underline{\mathbf{K}}}) \quad (117)$$

where $\dot{p} = \sqrt{2/3 \dot{\underline{\underline{\mathbf{E}}}}_h^p : \dot{\underline{\underline{\mathbf{E}}}}_h^p}$ is the cumulative plastic strain in the present context. These choices respectively provide the following regularisation partial differential equations:

$$E_{heq} = \chi - \text{Div} \rho_0 \frac{\partial \psi_0}{\partial \underline{\underline{\mathbf{K}}}}, \quad \text{or} \quad p = \chi - \text{Div} \rho_0 \frac{\partial \psi_0}{\partial \underline{\underline{\mathbf{K}}}} \quad (118)$$

If, $\underline{\underline{\mathbf{b}}}_0 = A \underline{\underline{\mathbf{K}}}$, then the regularisation operator involves the Lagrangian Laplace operator Δ_χ in the same way as in Eq. (97).

Finite Deformation Micromorphic Viscoplasticity Using Local Objective Frames

An alternative and frequently used method to formulate anisotropic elastoviscoplastic constitutive equations at finite deformations that identically fulfill the condition of Euclidean invariance (also called material frame indifference, see Besson et al. 2009), is to resort to local objective rotating frames, as initially proposed by Dogui and Sidoroff (1985) and Sidoroff and Dogui (2001). A local objective rotating frame is defined by the rotation field $\underline{\underline{\mathbf{Q}}}(\underline{\underline{\mathbf{x}}}, t)$, objective w.r.t. to further change of observer, and taking different values at different material points and different times:

$$\underline{\underline{\mathbf{x}}}^\dagger = \underline{\underline{\mathbf{Q}}}^T(\underline{\underline{\mathbf{x}}}, t) \cdot \underline{\underline{\mathbf{x}}} \quad (119)$$

It is based on the idea that there exists for each material point a privileged observer w.r.t. which the constitutive law takes a simple form. The method is described in details in Besson et al. (2009) and is used in many commercial finite element codes with the standard choices: corotational frame, such that $\underline{\underline{\mathbf{W}}}^\dagger = \underline{\underline{\dot{\mathbf{Q}}}} \cdot \underline{\underline{\mathbf{Q}}}^T = \underline{\underline{\mathbf{W}}}$, $\underline{\underline{\mathbf{W}}}$ being the skew-symmetric part of the velocity gradient,

and polar frame, such that $\underline{\underline{Q}}(\underline{x}, t) = \underline{\underline{R}}(\underline{x}, t)$, $\underline{\underline{R}}$ being the rotation part in the polar decomposition of the deformation gradient $\underline{\underline{F}}$. The main drawback of this method is the absence of thermodynamic background since, depending on specific constitutive choices within this framework, a free energy function of the strain may not exist.

This method is now applied to a micromorphic model including a scalar additional d.o.f. χ . Extension to higher order tensor-valued additional degrees of freedom is straightforward. The field equations are still given by (87). The stresses w.r.t. the local objective frames are:

$$\underline{\underline{\sigma}}^\dagger := \underline{\underline{Q}}^T \cdot \underline{\underline{\sigma}} \cdot \underline{\underline{Q}}, \quad a^\dagger = a, \quad \underline{\underline{b}}^\dagger = \underline{\underline{Q}}^T \cdot \underline{\underline{b}} \quad (120)$$

Time-derivation of these relations shows that the rotated stress derivatives are given by

$$\underline{\underline{Q}} \cdot \underline{\underline{\dot{\sigma}}}^\dagger \cdot \underline{\underline{Q}}^T = \underline{\underline{\dot{\sigma}}} + \underline{\underline{\sigma}} \cdot \underline{\underline{W}}^\dagger - \underline{\underline{W}}^\dagger \cdot \underline{\underline{\sigma}}, \quad \underline{\underline{Q}} \cdot \underline{\underline{\dot{b}}}^\dagger = \underline{\underline{\dot{b}}} - \underline{\underline{W}}^\dagger \cdot \underline{\underline{b}} \quad (121)$$

i.e. objective derivatives of the corresponding Eulerian stress tensors. If the corotational frame is used, the corresponding time derivative is the Jaumann rate, whereas it is the Green–Naghdi stress rate when the polar rotation is used. The same procedure is applied to the strain rates:

$$\underline{\underline{D}}^\dagger := \underline{\underline{Q}}^T \cdot \underline{\underline{D}} \cdot \underline{\underline{Q}}, \quad \underline{\underline{k}}^\dagger = \underline{\underline{Q}}^T \cdot \nabla_x \dot{\chi} \quad (122)$$

The time integration of the second equation in the rotating frame provides the variable $\underline{\underline{k}}^\dagger$. It must be underlined that $\underline{\underline{k}}^\dagger$ is NOT equal to $\underline{\underline{Q}}^T \cdot \nabla_x \dot{\chi}$. It is NOT the exact material time derivative of a constitutive variable, in general. The standard procedure then consists in postulating an additive decomposition of the rotated strain rate into elastic and plastic parts as

$$\underline{\underline{D}}^\dagger = \underline{\underline{\dot{e}}}^e + \underline{\underline{\dot{e}}}^p \quad (123)$$

where the elastic and plastic strain $\underline{\underline{e}}^e$ and $\underline{\underline{e}}^p$ are solely defined in the rotated frame. Anisotropic elastic laws are assumed to take the form:

$$\underline{\underline{\sigma}}^\dagger = \underline{\underline{c}} : \underline{\underline{e}}^e, \quad a = -H_\chi (p - \chi), \quad \underline{\underline{b}}^\dagger = \underline{\underline{A}} \cdot \underline{\underline{b}}^\dagger \quad (124)$$

Time-derivation of these equations and consideration of Eq. (120) show that the elasticity laws are in fact hypoelastic and that, generally, there does not exist a free energy density function from which they can be derived (Toll 2011).

The yield function and the flow rule are formulated within the rotated frame:

$$f(\tilde{\boldsymbol{\sigma}}^\dagger, R) = \sigma_{eq}^\dagger - R, \quad \dot{\tilde{e}}^p = \dot{p} \frac{\partial f}{\partial \tilde{\boldsymbol{\sigma}}^\dagger} \quad (125)$$

where normality is assumed for convenience and the viscoplastic multiplier \dot{p} is given by some viscoplastic law. The evolution of internal variables is of the form $\dot{q} = H(q, \tilde{e}^p)$ for suitable functions H . The yield radius is chosen as the following expression inspired by the previous thermodynamically based models: $R = R_{ref}(p) - a = R_{ref} - \text{div } \underline{\mathbf{b}}$. This extension of the micromorphic approach to finite deformations using rotating frames has been proposed first by Saanouni and Hamed (2013) and used by these authors for metal forming simulations involving regularised damage laws. As a result, the regularisation operator can be written as:

$$p = \chi - \frac{1}{H_\chi} \text{div } \underline{\mathbf{b}} = \chi - \frac{1}{H_\chi} \text{div} \left(\underline{\mathbf{Q}} \cdot \underline{\mathbf{b}}^\dagger \right) = \chi - \frac{1}{H_\chi} \text{div} \left(\underline{\mathbf{Q}} \cdot \underline{\mathbf{A}} \cdot \underline{\mathbf{k}}^\dagger \right) \quad (126)$$

In the isotropic case, $\underline{\mathbf{A}} = A \mathbf{1}$, the regularisation operator reduces to

$$p = \chi - \frac{A}{H_\chi} \text{div} \left(\underline{\mathbf{Q}} \cdot \underline{\mathbf{k}}^\dagger \right) \quad (127)$$

It is worth insisting on the fact that, in general, $\underline{\mathbf{Q}} \cdot \underline{\mathbf{k}}^\dagger \neq \nabla_x \chi$. Accordingly, the previous equation does not involve the Laplace operator and the regularisation is therefore nonlinear even with respect to rotated quantities.

Among all choices of rotating frames, the one associated with the logarithmic spin rate tensor (Xiao et al. 1999) was claimed to be the only one such that, when $\dot{\tilde{e}}^p = 0$, the isotropic hypoelastic strain-stress relation turns out to be hyperelastic. However, this property does not pertain to the general case $\dot{\tilde{e}}^p \neq 0$, so that this specific choice does not in general provide any explicit form of the regularisation operator.

Alternative constitutive choices are possible for the higher order stresses if Laplacian operators are preferred. They amount to restricting the use of the rotating frame only to the classical elastoviscoplasticity equations and to writing independently, $\underline{\mathbf{b}} = A \nabla_x \chi$, so that the regularisation operator is expressed in terms of the Eulerian Laplace operator Δ_x , see Eq. (148) in the next section, or $\underline{\mathbf{b}}_0 = A \nabla_X \chi$ which leads to the Lagrangian Laplace operator Δ_X , see Eq. (143) in the next section.

Note that limitations in the formulation of anisotropic plasticity arise from using symmetric plastic strain variable \tilde{e}^p and that generalisations are needed in order to introduce necessary plastic spins for materials with microstructures, which is possible within the rotating frame approach (Forest and Pilvin 1999).

Finite Deformation Micromorphic Elastoviscoplasticity Based on the Multiplicative Decomposition

The most appropriate thermodynamically based framework for the formulation of finite deformation anisotropic elastoviscoplasticity relies on the multiplicative decomposition of the deformation gradient, as settled by Mandel (1971). This method is applied here to a generalised continuum model again limited to one scalar degree of freedom, χ , in addition to the displacement vector, \underline{u} . The gradients of the degrees of freedom can be computed with respect to the reference or current coordinates:

$$\underline{\tilde{F}} = \underline{\tilde{1}} + \text{Grad } \underline{u} = \underline{\tilde{1}} + \underline{u} \otimes \nabla_X \quad (128)$$

$$\underline{\tilde{K}} = \text{Grad } \chi = \nabla_X \chi, \quad \underline{k} = \text{grad } \chi = \nabla_x \chi = \underline{\tilde{F}}^{-T} \cdot \underline{\tilde{K}} = \underline{\tilde{K}} \cdot \underline{\tilde{F}}^{-1} \quad (129)$$

The consideration of microdeformation degrees of freedom of higher order is possible without fundamental modification of the approach below, see Forest and Sievert (2006).

A multiplicative decomposition is envisaged in this section, see Eq. (12), in the form:

$$\underline{\tilde{F}} = \underline{\tilde{F}}^e \cdot \underline{\tilde{F}}^p \quad (130)$$

which assumes the existence of a triad of directors attached to the material point in order to unambiguously define the isoclinic intermediate local configuration, labelled (#) in the sequel, see Mandel (1971, 1973). The directors are usually related to non-material microstructure directions like lattice directions in single crystals or fibre directions in composites. The existence of such directors is required for the formulation of objective anisotropic constitutive equations (Besson et al. 2009). The Jacobians of all contributions in Eq. (130) are denoted by

$$J = \det \underline{\tilde{F}}, \quad J_e = \det \underline{\tilde{F}}^e, \quad J_p = \det \underline{\tilde{F}}^p, \quad J = J_e J_p, \quad \rho_0 = \rho_{\#} J_p = \rho J \quad (131)$$

They are used to relate the mass densities with respect to the three local configurations: In the present section, the microdeformation gradient $\underline{\tilde{K}}$ is not split into elastic and plastic contributions, although it is possible as done in section “Elastic-Plastic Decomposition of the Generalized Strain Measures,” at the expense of additional evolution laws to be determined and of drastically different regularisation operators.

The power density of internal and contact forces are

$$p^{(i)} = \underline{\tilde{\sigma}} : \dot{\underline{\tilde{F}}} \underline{\tilde{F}}^{-1} + a \dot{\chi} + \underline{k} \cdot \nabla_x \dot{\chi}, \quad \forall \underline{x} \in \Omega, \quad p^{(c)} = \underline{t} \cdot \underline{u} + a_c \dot{\chi}, \quad \forall \underline{x} \in \partial\Omega \quad (132)$$

The invariance of $p^{(i)}$ with respect to any change of observer requires the Cauchy stress $\underline{\sigma}$ to be symmetric. The scalar microstrain is assumed to be invariant. The corresponding balance and boundary conditions are still given by Eq. (87).

Lagrangian Formulation

The Lagrangian free energy density is a function $\psi_0(\underline{\mathbf{C}}^e, q, \chi, \underline{\mathbf{K}})$, where $\underline{\mathbf{C}}^e = \underline{\mathbf{F}}^{eT} \cdot \underline{\mathbf{F}}^e$ is the elastic strain and q a set of internal variables accounting for material hardening. Note that the usual elastic strain tensor $\underline{\mathbf{C}}^e$ is defined with respect to the intermediate configuration to comply with standard anisotropic plasticity, whereas $\underline{\mathbf{K}}$ is Lagrangian. The presented formulation is therefore not purely Lagrangian but rather mixed. The local Lagrangian form of the entropy inequality is: $\mathcal{D}_0 = p_0^{(i)} - p_0 \dot{\psi}_0 \geq 0$, $p_0^{(i)} = J p^{(i)}$. Accounting for the multiplicative decomposition (130), the power of internal forces is expanded as:

$$\begin{aligned} p_0^{(i)} &= J \underline{\underline{\sigma}} : \underline{\underline{\dot{\mathbf{F}}}} \cdot \underline{\underline{\mathbf{F}}}^{-1} + J a \dot{\chi} + J \underline{\underline{\mathbf{b}}} \cdot \nabla_x \dot{\chi} = \frac{J_p}{2} \underline{\underline{\Pi}}^e : \underline{\underline{\dot{\mathbf{C}}}}^e \\ &+ J_p \underline{\underline{\Pi}}^M : \underline{\underline{\dot{\mathbf{F}}}}^p \cdot \underline{\underline{\mathbf{F}}}^{p-1} + a_0 \dot{\chi} + \underline{\underline{\mathbf{b}}}_0 \cdot \underline{\underline{\dot{\mathbf{K}}}} \end{aligned} \quad (133)$$

where the Piola stress tensor w.r.t. the intermediate configuration and the Mandel stress tensor according to (Haupt 2000) are respectively defined as:

$$\underline{\underline{\Pi}}^e = J_e \underline{\underline{\mathbf{F}}}^{e-1} \cdot \underline{\underline{\sigma}} \cdot \underline{\underline{\mathbf{F}}}^{e-T}, \quad \underline{\underline{\Pi}}^M = \underline{\underline{\mathbf{C}}}^e \cdot \underline{\underline{\Pi}}^e = J_e \underline{\underline{\mathbf{F}}}^{eT} \cdot \underline{\underline{\sigma}} \cdot \underline{\underline{\mathbf{F}}}^{e-T} \quad (134)$$

The Lagrangian generalized stresses in (133) are $a_0 = Ja$ and $\underline{\underline{\mathbf{b}}}_0 = J \underline{\underline{\mathbf{F}}}^{-1} \underline{\underline{\mathbf{b}}}$. As a result the dissipation rate becomes:

$$\begin{aligned} &\left(\frac{J_p}{2} \underline{\underline{\Pi}}^e - \rho_0 \frac{\partial \psi_0}{\partial \underline{\underline{\mathbf{C}}}} \right) : \underline{\underline{\dot{\mathbf{C}}}}^e + \left(a_0 - \rho_0 \frac{\partial \psi_0}{\partial \chi} \right) \dot{\chi} + \left(\underline{\underline{\mathbf{b}}}_0 - \rho_0 \frac{\partial \psi_0}{\partial \underline{\underline{\mathbf{K}}}} \right) \cdot \underline{\underline{\dot{\mathbf{K}}}} \\ &+ J_p \underline{\underline{\Pi}}^M : \underline{\underline{\dot{\mathbf{F}}}}^p \cdot \underline{\underline{\mathbf{F}}}^{p-1} - \rho_0 \frac{\partial \psi_0}{\partial q} \dot{q} \geq 0 \end{aligned} \quad (135)$$

The following state and evolution laws ensure the positivity of \mathcal{D}_0 :

$$\underline{\underline{\Pi}}^e = 2\rho_0 \frac{\partial \psi_0}{\partial \underline{\underline{\mathbf{C}}}}, \quad a_0 = \rho_0 \frac{\partial \psi_0}{\partial \chi}, \quad \underline{\underline{\mathbf{b}}}_0 = \rho_0 \frac{\partial \psi_0}{\partial \underline{\underline{\mathbf{K}}}}, \quad R = \rho_0 \frac{\partial \psi_0}{\partial q} \quad (136)$$

$$\underline{\underline{\dot{\mathbf{F}}}}^p \cdot \underline{\underline{\dot{\mathbf{F}}}}^{p-1} = \frac{\partial \Omega}{\partial \underline{\underline{\Pi}}^M}(\underline{\underline{\Pi}}^M, R), \quad \dot{q} = -\frac{\partial \Omega}{\partial R}(\underline{\underline{\Pi}}^M, R), \quad (137)$$

Following Mandel (1971), a dissipation potential $\Omega(\tilde{\mathbf{\Pi}}^M, R)$ function of the driving forces for plasticity, is introduced to formulate the flow and hardening variable evolution rule. If the dissipation function is convex w.r.t. $\tilde{\mathbf{\Pi}}^M$ and concave w.r.t. R , the positivity of dissipation is ensured. Specific expressions for within the context of viscoplasticity can be found in Besson et al. (2009).

As an example, the following free energy potential is proposed:

$$\rho_0 \psi_0 = \frac{1}{2} J_p \tilde{\mathbf{E}}^e : \tilde{\mathbf{c}} : \tilde{\mathbf{E}}^e + \rho_0 \psi_q(q) + \frac{1}{2} H_\chi (p - \chi)^2 + \rho_0 \psi_{\nabla_X}(\underline{\mathbf{K}}) \quad (138)$$

where ψ_q is the appropriate free energy contribution associated with usual work-hardening (not specified here) and $2\tilde{\mathbf{E}}^e = \tilde{\mathbf{C}}^e - \tilde{\mathbf{1}}$ is the Green-Lagrange elastic strain measure. The microstrain variable is compared to the cumulative plastic strain variable p defined as:

$$\dot{p} = \sqrt{\frac{2}{3} \left(\dot{\tilde{\mathbf{P}}} \cdot \tilde{\mathbf{P}}^{-1} \right) : \left(\dot{\tilde{\mathbf{P}}} \cdot \tilde{\mathbf{P}}^{-1} \right)} \quad (139)$$

According to the state laws (136), we obtain

$$\tilde{\mathbf{\Pi}}^e = \rho_0 \frac{\partial \psi_0}{\partial \tilde{\mathbf{E}}^e} = \tilde{\mathbf{c}} : \tilde{\mathbf{E}}^e, \quad a_0 = -H_\chi (p - \chi), \quad \underline{\mathbf{b}}_0 = \rho_0 \frac{\partial \psi_{\nabla_X}}{\partial \underline{\mathbf{K}}} \quad (140)$$

The regularisation operator then follows from the combination of the previous constitutive equations with the balance equation (89):

$$p = \chi - \frac{1}{H_\chi} \text{Div} \left(\rho_0 \frac{\partial \psi_{\nabla_X}}{\partial \underline{\mathbf{K}}} \right) \quad (141)$$

The specific choice $\rho_0 \psi_{\nabla_X}(\underline{\mathbf{K}}) = \frac{1}{2} \underline{\mathbf{K}} \cdot \tilde{\mathbf{A}} \cdot \underline{\mathbf{K}}$ leads to a regularisation operator that is linear with respect to Lagrangian coordinates:

$$p = \chi - \frac{1}{H_\chi} \text{Div} \left(\tilde{\mathbf{A}} \cdot \text{Grad } \chi \right) \quad (142)$$

which involves the Laplacian operator Δ_X in the isotropic case, i.e. $\tilde{\mathbf{A}} = A \mathbf{1}$,

$$\text{Op} = 1 - \frac{A}{H_\chi} \Delta_X \quad (143)$$

The impact on hardening can be seen by choosing, as an example, $q = p$, according to (139), as an internal variable in (138). The dissipation potential can be chosen in such a way that the residual dissipation takes the form

$$\mathcal{D}_0 = J_p \underline{\Pi}^M : \left(\dot{\underline{\mathbf{F}}}^p \cdot \underline{\mathbf{F}}^{p-1} \right) - \rho_0 \frac{\partial \Psi_q}{\partial p} \dot{p} = J_p \Pi_{eq}^M \dot{p} - R \dot{p} \geq 0 \quad (144)$$

with

$$\dot{p} = \frac{\partial \Omega}{\partial f}, \quad f \left(\underline{\Pi}^M, R \right) = J_p \Pi_{eq}^M - R$$

where f is the yield function. As a result, the yield stress R is given by the following enhanced hardening law:

$$R = \rho_0 \frac{\partial \psi_0}{\partial p} = \rho_0 \frac{\partial \psi_q}{\partial p} + H_\chi (p - \chi) = \rho_0 \frac{\partial \psi_q}{\partial p} - \text{Div} \left(\rho_0 \frac{\partial \psi_{\nabla \chi}}{\partial \underline{\mathbf{K}}} \right) \quad (145)$$

Eulerian Formulation

In the Eulerian formulation, the free energy density is taken as a function of $\underline{\mathbf{k}}$ instead of $\underline{\mathbf{K}}$, according to (129): $\psi \left(\underline{\mathbf{C}}^e, \chi, \underline{\mathbf{k}} \right)$, so that the state laws for generalised stresses become:

$$\underline{\mathbf{a}} = \rho \frac{\partial \psi}{\partial \chi}, \quad \underline{\mathbf{b}} = \rho \frac{\partial \psi}{\partial \underline{\mathbf{k}}} \quad (146)$$

The arguments of the free energy mix the invariant quantities $\underline{\mathbf{C}}^e$, χ and the observer-dependent variable $\underline{\mathbf{k}}$. Galilean invariance then requires ψ to be isotropic with respect to $\underline{\mathbf{k}}$.

The constitutive choice (138) is now replaced by

$$\rho \psi = \frac{1}{2} J \underline{\mathbf{E}}^e : \underline{\underline{\mathbf{c}}} : \underline{\underline{\mathbf{E}}}^e + \rho \psi_q(q) + \frac{1}{2} H_\chi (p - \chi)^2 + \rho \psi_\nabla(\underline{\mathbf{k}}) \quad (147)$$

A quadratic potential ψ_∇ is necessarily of the form $A \|\underline{\mathbf{k}}\|^2/2$, for objectivity reasons, so that the regularisation operator involves the Laplace operator Δ_x w.r.t. Eulerian coordinates:

$$\text{Op} = 1 - \frac{A}{H_\chi} \Delta_x \quad (148)$$

If the same viscoplastic yield function $f \left(\underline{\Pi}^M, R \right)$ as in the previous subsection is adopted, the hardening rule is enhanced as follows:

$$R = \rho \frac{\partial \psi}{\partial p} = \rho \frac{\partial \psi_q}{\partial p} + H_\chi (p - \chi) = \rho \frac{\partial \psi_q}{\partial p} - \text{div} \underline{\mathbf{b}} = \rho \frac{\partial \psi_q}{\partial p} - A \Delta_x \chi \quad (149)$$

This therefore yields a finite strain generalisation of Aifantis strain gradient plasticity model (Aifantis 1987; Forest and Aifantis 2010).

Formulation Using the Local Intermediate Configuration Only

In the two previous formulations, Lagrangian or Eulerian generalized strain variables were mixed with the elastic strain variable $\underline{\underline{C}}^e$ attached to the intermediate local configuration, as the arguments of the free energy function. It is possible to assign the free energy function with a consistent set of arguments solely attached to the intermediate configuration. For that purpose, a generalised strain $\underline{\underline{K}}^\sharp$ and generalised stresses $a_\sharp, \underline{\underline{b}}^\sharp$ are now defined on the intermediate local configuration:

$$\underline{\underline{K}}^\sharp = \underline{\underline{k}} \cdot \underline{\underline{F}}^e = \underline{\underline{F}}^{eT} \cdot \underline{\underline{k}} = \underline{\underline{K}} \cdot \underline{\underline{F}}^{p-1} = \underline{\underline{F}}^{p-T} \cdot \underline{\underline{K}} \quad (150)$$

$$a_\sharp = J_e a = J_p^{-1} a_0, \quad \underline{\underline{b}}^\sharp = J_p^{-1} \underline{\underline{F}}^p \cdot \underline{\underline{b}}_0 = J_e \underline{\underline{F}}^{e-1} \cdot \underline{\underline{b}} \quad (151)$$

The power density of internal forces expressed w.r.t. the intermediate local configuration then takes the form:

$$p_\sharp^{(i)} = J_p^{-1} p_0^{(i)} = \frac{1}{2} \underline{\underline{\Pi}}^e : \underline{\underline{\dot{C}}}^e + \left(\underline{\underline{\Pi}}^M + \underline{\underline{K}}^\sharp \otimes \underline{\underline{b}}^\sharp \right) : \underline{\underline{\dot{F}}}^p \cdot \underline{\underline{F}}^{p-1} + a_\sharp \dot{\chi} + \underline{\underline{b}}^\sharp \cdot \underline{\underline{\dot{K}}}^\sharp \quad (152)$$

where $p_0^{(i)}$ is still given by Eq. (133). To establish this expression, the following relation was used:

$$\underline{\underline{\dot{K}}}^\sharp = \underline{\underline{F}}^{pT} \cdot \underline{\underline{\dot{K}}}^\sharp + \underline{\underline{\dot{F}}}^{pT} \cdot \underline{\underline{K}}^\sharp \quad (153)$$

The dissipation rate density measured w.r.t. the intermediate local configuration is then:

$$\mathcal{D}_\sharp = p_\sharp^{(i)} - \rho_\sharp \dot{\psi}_\sharp \geq 0 \quad (154)$$

The free energy density function is chosen as $\psi_\sharp \left(\underline{\underline{C}}^e, q, \chi, \underline{\underline{K}}^\sharp \right)$. As such, it is invariant w.r.t. change of observer. The Clausius–Duhem inequality is now derived as

$$\begin{aligned} & \left(\frac{1}{2} \underline{\underline{\Pi}}^e - \rho_\sharp \frac{\partial \psi_\sharp}{\partial \underline{\underline{C}}^e} \right) : \underline{\underline{\dot{C}}}^e + \left(a_\sharp - \rho_\sharp \frac{\partial \psi_\sharp}{\partial \chi} \right) \dot{\chi} + \left(\underline{\underline{b}}^\sharp - \rho_\sharp \frac{\partial \psi_\sharp}{\partial \underline{\underline{K}}^\sharp} \right) \cdot \underline{\underline{\dot{K}}}^\sharp \\ & + \left(\underline{\underline{\Pi}}^M + \underline{\underline{K}}^\sharp \otimes \underline{\underline{b}}^\sharp \right) : \underline{\underline{\dot{F}}}^p \cdot \underline{\underline{F}}^{p-1} - \rho_\sharp \frac{\partial \psi_\sharp}{\partial q} \dot{q} \geq 0 \end{aligned} \quad (155)$$

This expression reveals the existence of a generalised Mandel tensor, $\tilde{\Pi}^M + \underline{\mathbf{K}}^\sharp \otimes \underline{\mathbf{b}}^\sharp$, conjugate to the plastic deformation rate, that is a function of the classical Mandel stress and of microdeformation related stress and strain. Positivity of dissipation is ensured by the choice of the following state laws and plastic flow and hardening rules:

$$\tilde{\Pi}^e = 2\rho_\sharp \frac{\partial \psi_\sharp}{\partial \tilde{\mathbf{C}}^e}, \quad a_\sharp = \rho_\sharp \frac{\partial \psi_\sharp}{\partial \chi}, \quad \underline{\mathbf{b}}^\sharp = \rho_\sharp \frac{\partial \psi_\sharp}{\partial \underline{\mathbf{K}}^\sharp}, \quad R = \rho_\sharp \frac{\partial \psi_\sharp}{\partial q} \quad (156)$$

$$\dot{\tilde{\mathbf{F}}}^p \cdot \tilde{\mathbf{F}}^{p-1} = \frac{\partial \Omega}{\partial (\tilde{\Pi}^M + \underline{\mathbf{K}}^\sharp \otimes \underline{\mathbf{b}}^\sharp)}, \quad \dot{q} = -\frac{\partial \Omega}{\partial R} \quad (157)$$

provided that the dissipation potential $\Omega(\tilde{\Pi}^M + \underline{\mathbf{K}}^\sharp \otimes \underline{\mathbf{b}}^\sharp, R)$ displays suitable convexity properties with respect to both arguments.

The yield criterion is taken as a function $f(\tilde{\Pi}^M + \underline{\mathbf{K}}^\sharp \otimes \underline{\mathbf{b}}^\sharp, R) = \Pi_{eq}^M - R$ where the Π_{eq}^M is an equivalent stress measure based on the generalized Mandel stress tensor. Choosing $q = p$, where p is still given by Eq. (139), the residual dissipation takes the form:

$$\mathcal{D}_\sharp = (\tilde{\Pi}^M + \underline{\mathbf{K}}^\sharp \otimes \underline{\mathbf{b}}^\sharp) : \dot{\tilde{\mathbf{F}}}^p \cdot \tilde{\mathbf{F}}^{p-1} - R\dot{p} = \Pi_{eq}^M \dot{p} - R\dot{p} \quad (158)$$

Note that the contribution $\underline{\mathbf{K}}^\sharp \otimes \underline{\mathbf{b}}^\sharp$ in the generalised Mandel stress acts as a size-dependent kinematic hardening component which comes in addition to isotropic hardening represented by R . This is a specific feature of the model formulation w.r.t. the intermediate configuration.

As an example, a typical form of the free energy density function based on constitutive variables defined on the intermediate configuration, and hyperelastic laws are:

$$\rho_\sharp \psi_\sharp = \frac{1}{2} \tilde{\mathbf{E}}^e : \tilde{\mathbf{c}} : \tilde{\mathbf{E}}^e + \rho_\sharp \psi_\sharp(q) + \frac{1}{2} H_\chi (p - \chi)^2 + \frac{1}{2} \rho_\sharp \psi_{\sharp\nabla}(\underline{\mathbf{K}}^\sharp) \quad (159)$$

$$\tilde{\Pi}^e = \tilde{\mathbf{c}} : \tilde{\mathbf{E}}^e, \quad a_\sharp = -H_\chi (p - \chi), \quad \underline{\mathbf{b}}^\sharp = \rho_\sharp \frac{\partial \psi_\sharp}{\partial \underline{\mathbf{K}}^\sharp} \quad (160)$$

These generalized stresses can be inserted into the balance equation

$$a_0 = \text{Div } \underline{\mathbf{b}}_0 \Rightarrow J_p a_\sharp = \text{Div} \left(J_p \tilde{\mathbf{F}}^{p-1} \cdot \underline{\mathbf{b}}^\sharp \right) \quad (161)$$

This provides the form of the regularisation operator:

$$p = \chi - \frac{1}{J_p H_\chi} \text{Div} \left(J_p \underline{\mathbf{F}}^{p-1} \rho_\# \frac{\partial \psi_\#}{\partial \underline{\mathbf{K}}^\#} \right) \quad (162)$$

A quadratic dependence of the contribution $\rho_\# \psi_\# \nabla = 1/2 \underline{\mathbf{K}}^\# \cdot \underline{\mathbf{A}} \cdot \underline{\mathbf{K}}^\#$ leads to the following linear relationship between $\underline{\mathbf{b}}^\#$ and $\underline{\mathbf{K}}^\#$:

$$\underline{\mathbf{b}}^\# = A \underline{\mathbf{K}}^\# \Rightarrow \underline{\mathbf{b}} = J_e^{-1} A \underline{\mathbf{B}}^e \cdot \underline{\mathbf{k}} \quad \text{and} \quad \underline{\mathbf{b}}^\# = J_p A \underline{\mathbf{C}}^{p-1} \cdot \underline{\mathbf{k}} \quad (163)$$

with $\underline{\mathbf{B}}^e = \underline{\mathbf{F}}^e \cdot \underline{\mathbf{F}}^{eT}$ and $\underline{\mathbf{C}}^p = \underline{\mathbf{F}}^{pT} \cdot \underline{\mathbf{F}}^p$. However, in that case, the regularisation operator (161) is nonlinear and does not involve a Laplace operator, even in the isotropic case $\underline{\mathbf{A}} = A \underline{\mathbf{1}}$. As a result, the hyperelastic relationships for the higher order stresses are not linear w.r.t. to the associated strain gradient measures.

Conclusion

Eringen and Mindlin's micromorphic theory offers real opportunities for the modeling of size effects in the mechanics of materials. Elastic-viscoplastic constitutive laws have been formulated at finite deformations. They remain to be further specialized and calibrated with respect to size effects observed in metal and polymer plasticity. Successful applications deal for example with the ductile fracture of metallic alloys (Enakoutsa and Leblond 2009; Hutter 2017b) and porous metals (Dillard et al. 2006). Micromorphic elasticity has been recently revisited and further developed to account for the dispersion of elastic waves in architected and metamaterials (Neff et al. 2014; Rosi and Auffray 2016; Madeo et al. 2016). Intensive work is needed to establish connections between the micromorphic continuum theories and the actual underlying microstructure (Forest and Trinh 2011; Hutter 2017a; Biswas and Poh 2017).

The proposed systematic treatment of the thermomechanics of continua with additional degrees of freedom leads to model formulations ranging from micromorphic to phase field models. In particular, a general framework for the introduction of dissipative processes associated with the additional degrees of freedom has been proposed. If internal constraints are enforced on the relation between macro and microvariables in the micromorphic approach, standard second gradient and strain gradient plasticity models can be retrieved.

As a variant of micromorphic continuum, microdamage continuum and its regularization capabilities for the modelling of crack propagation in single crystals have been studied. First, a crystallographic constitutive model which accounts for continuum damage with respect to fracture planes has been presented. Then, the theory has been extended from classical continuum to microdamage continuum. It has been shown that the approach can be a good candidate for solving mesh

dependency and the prediction of final fracture in anisotropic media. Analytical fits and numerical results showed that the theory is well suited for FEA and possesses a great potential for the future modelling aspects. Comparison with available data on crack growth especially cyclic loading in nickel-based superalloys, will be decisive to conclude on the ability of the approach to reach realistic prediction of component failure.

The presented extensions to finite deformations show that the regularisation operator cannot be postulated in an intuitive way. It is rather the result of a constitutive choice regarding the dependence of the free energy function on the gradient term. Purely Lagrangian and Eulerian formulations are straightforward and lead to Helmholtz-like operators w.r.t. Lagrangian or Eulerian coordinates. Two alternative standard procedures of extension of classical constitutive laws to large strains, widely used in commercial finite element codes, have been combined with the micromorphic approach. In particular, the choice of local objective rotating frames leads to new nonlinear regularisation operators that are not of the Helmholtz type. Three distinct operators were proposed within the context of the multiplicative decomposition of the deformation gradient. A new feature is that a free energy density function depending on variables solely defined with respect to the intermediate isoclinic configuration leads to the existence of additional kinematic hardening induced by the gradient of a scalar micromorphic degree of freedom.

Note that the results obtained for the micromorphic theory with additional degrees of freedom are also valid for gradient theories (gradient plasticity or gradient damage) once an internal constraint is imposed linking the additional degrees of freedom to strain or internal variables. This amounts to selecting high values of parameter H_χ or introducing corresponding Lagrange multipliers. The analysis was essentially limited to scalar micromorphic degrees of freedom for the sake of simplicity, even though tensorial examples were also given. Scalar plastic microstrain approaches suffer from limitations like indeterminacy of flow direction at cusps of the cumulative plastic strain in bending for instance, see Peerlings (2007), Poh et al. (2011), and Wulfinghoff et al. (2014). Those limitations can be removed by the use of tensorial micromorphic variable (microstrain or microdeformation tensors). The micromorphic approach is not limited to the gradient of strain-like, damage or phase field variables. It can also be applied to other internal variables, as demonstrated for hardening variables in Dorgan and Voyiadjis (2003) and Saanouni and Hamed (2013).

It remains that the regularisation properties of the derived nonlinear operators are essentially unknown, except through examples existing in the mentioned literature. For instance, the Eulerian and Lagrangian variants of the Helmholtz-type equation for scalar micromorphic strain variables have been assessed by Wcislo et al. (2013) giving the advantage to the latter, based on finite element simulations of specific situations. The regularising properties of more general operators should be explored in the future from the mathematical and computational perspectives in order to select the most relevant constitutive choices that may depend on the type of material classes.

It may be surprising that the constitutive theory underlying the construction of regularisation operators for plasticity and damage, mainly relies on the enhancement of the free energy density function instead of the dissipative laws. It is in fact widely recognised that plastic strain gradients, e.g. associated with the multiplication of geometrically necessary dislocations, lead to energy storage that can be released by further deformation or heat treatments. However, the limitation to the enhancement of free energy potential is mainly due to the simplicity of the theoretical treatment and to the computational efficiency of the operators derived in that way. Dissipative higher order contributions remain to be explored from the viewpoint of regularisation, as started in Forest (2009). Constitutive models of that kind are already available for plasticity, damage and fracture (Amor et al. 2009; Pham et al. 2011; Vignollet et al. 2014; Miehe et al. 2016).

Acknowledgments The first author thanks Prof. O. Aslan for his contribution to the presented micro-morphic damage theory. These contributions are duly cited in the references quoted in the text and listed below.

References

- E. Aifantis, On the microstructural origin of certain inelastic models. *J. Eng. Mater. Technol.* **106**, 326–330 (1984)
- E. Aifantis, The physics of plastic deformation. *Int. J. Plast.* **3**, 211–248 (1987)
- K. Ammar, B. Appolaire, G. Cailletaud, F. Feyel, F. Forest, Finite element formulation of a phase field model based on the concept of generalized stresses. *Comput. Mater. Sci.* **45**, 800–805 (2009)
- H. Amor, J.J. Marigo, C. Maurini, Regularized formulation of the variational brittle fracture with unilateral contact: Numerical experiments. *J. Mech. Phys. Solids* **57**, 1209–1229 (2009)
- O. Aslan, S. Forest, Crack growth modelling in single crystals based on higher order continua. *Comput. Mater. Sci.* **45**, 756–761 (2009)
- O. Aslan, S. Forest, The micromorphic versus phase field approach to gradient plasticity and damage with application to cracking in metal single crystals, in *Multiscale Methods in Computational Mechanics*, ed. by R. de Borst, E. Ramm. Lecture Notes in Applied and Computational Mechanics, vol. 55, (Springer, New York, 2011), pp. 135–154
- O. Aslan, N.M. Cordero, A. Gaubert, S. Forest, Micromorphic approach to single crystal plasticity and damage. *Int. J. Eng. Sci.* **49**, 1311–1325 (2011)
- N. Auffray, H. Le Quang, Q. He, Matrix representations for 3D strain-gradient elasticity. *J. Mech. Phys. Solids* **61**, 1202–1223 (2013)
- A. Berezovski, J. Engelbrecht, P. Van, Weakly nonlocal thermoelasticity for microstructured solids: Microdeformation and microtemperature. *Arch. Appl. Mech.* **84**, 1249–1261 (2014)
- A. Bertram, *Elasticity and Plasticity of Large Deformations* (Springer, Heidelberg, 2012)
- Besson, J., Cailletaud, G., Chaboche, J.-L., Forest, S., Blety, M., *Non-linear Mechanics of Materials*. Solid Mechanics and Its Applications, vol. 167 (Springer, Dordrecht, 2009), 433 p. ISBN:978-90-481-3355-0
- R. Biswas, L. Poh, A micromorphic computational homogenization framework for heterogeneous materials. *J. Mech. Phys. Solids* **102**, 187–208 (2017)
- G. Cailletaud, S. Forest, D. Jeulin, F. Feyel, I. Galliet, V. Mounoury, S. Quilici, Some elements of microstructural mechanics. *Comput. Mater. Sci.* **27**, 351–374 (2003)
- A. L. Cauchy, Recherches sur l'équilibre et le mouvement intérieur des corps solides ou fluides, élastiques ou non élastiques. *Bulletin de la Société Philomatique*, 9–13 (1822)

- B. Coleman, W. Noll, The thermodynamics of elastic materials with heat conduction and viscosity. *Arch. Ration. Mech. Anal.* **13**, 167–178 (1963)
- T. Dillard, S. Forest, P. Ienny, Micromorphic continuum modelling of the deformation and fracture behaviour of nickel foams. *Eur. J. Mech. A/Solids* **25**, 526–549 (2006)
- A. Dogui, F. Sidoroff, Kinematic hardening in large elastoplastic strain. *Eng. Fract. Mech.* **21**, 685–695 (1985)
- J. Dorgan, G. Voyiadjis, Nonlocal dislocation based plasticity incorporating gradients of hardening. *Mech. Mater.* **35**, 721–732 (2003)
- K. Enakoutsa, J. Leblond, Numerical implementation and assessment of the glpd micromorphic model of ductile rupture. *Eur. J. Mech. A/Solids* **28**, 445–460 (2009)
- R. Engelen, M. Geers, F. Baaijens, Nonlocal implicit gradient-enhanced elasto-plasticity for the modelling of softening behaviour. *Int. J. Plast.* **19**, 403–433 (2003)
- A. Eringen, Theory of thermo-microstretch elastic solids. *Int. J. Eng. Sci.* **28**, 1291–1301 (1990)
- A. Eringen, *Microcontinuum Field Theories* (Springer, New York, 1999)
- A. Eringen, E. Suhubi, Nonlinear theory of simple microelastic solids. *Int. J. Eng. Sci.* **2**(189–203), 389–404 (1964)
- S. Forest, The micromorphic approach for gradient elasticity, viscoplasticity and damage. *ASCE J. Eng. Mech.* **135**, 117–131 (2009)
- S. Forest, Micromorphic media, in *Generalized Continua – From the Theory to Engineering Applications*. *CISM International Centre for Mechanical Sciences*, ed. by H. Altenbach, V. Eremeyev. Courses and Lectures No. 541 (Springer, New York, 2012), pp. 249–300
- S. Forest, Nonlinear regularisation operators as derived from the micromorphic approach to gradient elasticity, viscoplasticity and damage. *Proc. R. Soc. A* **472**, 20150755 (2016)
- S. Forest, E.C. Aifantis, Some links between recent gradient thermo-elasto-plasticity theories and the thermomechanics of generalized continua. *Int. J. Solids Struct.* **47**, 3367–3376 (2010)
- S. Forest, A. Bertram, Formulations of strain gradient plasticity, in *Mechanics of Generalized Continua*, ed. by H. Altenbach, G.A. Maugin, V. Erofeev. *Advanced Structured Materials*, vol. 7 (Springer, Berlin/Heidelberg, 2011), pp. 137–150
- S. Forest, P. Pilvin, Modelling finite deformation of polycrystals using local objective frames. *Z. Angew. Math. Mech.* **79**, S199–S202 (1999)
- S. Forest, K. Sab, Finite deformation second order micromorphic theory and its relations to strain and stress gradient models. *Math. Mech. Solids* (2017). <https://doi.org/10.1177/1081286517720844>
- S. Forest, R. Sievert, Elastoviscoplastic constitutive frameworks for generalized continua. *Acta Mech.* **160**, 71–111 (2003)
- S. Forest, R. Sievert, Nonlinear microstrain theories. *Int. J. Solids Struct.* **43**, 7224–7245 (2006)
- S. Forest, D.K. Trinh, Generalized continua and non-homogeneous boundary conditions in homogenization methods. *ZAMM* **91**, 90–109 (2011)
- M.G.D. Geers, Finite strain logarithmic hyperelasto-plasticity with softening: A strongly non-local implicit gradient framework. *Comput. Methods Appl. Mech. Eng.* **193**, 3377–3401 (2004)
- M. Geers, R.D. Borst, W. Brekelmans, R. Peerlings, On the use of local strain fields for the determination of the intrinsic length scale. *Journal de Physique IV* **8**(Pr8), 167–174 (1998)
- P. Germain, La méthode des puissances virtuelles en mécanique des milieux continus, première partie : théorie du second gradient. *J. de Mécanique* **12**, 235–274 (1973a)
- P. Germain, The method of virtual power in continuum mechanics. Part 2: microstructure. *SIAM J. Appl. Math.* **25**, 556–575 (1973b)
- P. Germain, Q. Nguyen, P. Suquet, Continuum thermodynamics. *J. Appl. Mech.* **50**, 1010–1020 (1983)
- N. Germain, J. Besson, F. Feyel, Simulation of laminate composites degradation using mesoscopic non-local damage model and non-local layered shell element. *Model. Simul. Mater. Sci. Eng.* **15**, S425–S434 (2007)
- M. Goodman, S. Cowin, A continuum theory for granular materials. *Arch. Ration. Mech. Anal.* **44**, 249–266 (1972)

- P. Grammenoudis, C. Tsakmakis, Micromorphic continuum part I: Strain and stress tensors and their associated rates. *Int. J. Non-Linear Mech.* **44**, 943–956 (2009)
- P. Grammenoudis, C. Tsakmakis, D. Hofer, Micromorphic continuum part II: Finite deformation plasticity coupled with damage. *Int. J. Non-Linear Mech.* **44**, 957–974 (2009)
- M. Gurtin, Generalized Ginzburg–Landau and Cahn–Hilliard equations based on a microforce balance. *Physica D* **92**, 178–192 (1996)
- P. Haupt, *Continuum Mechanics and Theory of Materials* (Springer, Berlin, 2000)
- T. Helfer, Extension of monodimensional fuel performance codes to finite strain analysis using a Lagrangian logarithmic strain framework. *Nucl. Eng. Des.* **288**, 75–81 (2015)
- C. Hirschberger, P. Steinmann, Classification of concepts in thermodynamically consistent generalized plasticity. *ASCE J. Eng. Mech.* **135**, 156–170 (2009)
- C. Hirschberger, E. Kuhl, P. Steinmann, On deformational and configurational mechanics of micromorphic hyperelasticity - theory and computation. *Comput. Methods Appl. Mech. Eng.* **196**, 4027–4044 (2007)
- M. Horak, M. Jirasek, An extension of small-strain models to the large-strain range based on an additive decomposition of a logarithmic strain. *Programs Algorithms Numer. Math.* **16**, 88–93 (2013)
- G. Hütter, Homogenization of a Cauchy continuum towards a micromorphic continuum. *J. Mech. Phys. Solids* **99**, 394–408 (2017a)
- G. Hütter, A micromechanical gradient extension of Gurson’s model of ductile damage within the theory of microdilatational media. *Int. J. Solids Struct.* **110–111**, 15–23 (2017b)
- C. Kafadar, A. Eringen, Micropolar media: I the classical theory. *Int. J. Eng. Sci.* **9**, 271–305 (1971)
- N. Kirchner, P. Steinmann, A unifying treatise on variational principles for gradient and micromorphic continua. *Philos. Mag.* **85**, 3875–3895 (2005)
- M. Lazar, G. Maugin, On microcontinuum field theories: The Eshelby stress tensor and incompatibility conditions. *Philos. Mag.* **87**, 3853–3870 (2007)
- E.H. Lee, P. Germain, Elastic–plastic theory at finite strain, in *Problems of Plasticity*, ed. by A. Sawczuk (Ed), (Noordhoff International Publishing, 1972), pp. 117–133
- A. Madeo, G. Barbagallo, M.V. d’Agostino, L. Placidi, P. Neff, First evidence of non-locality in real band-gap metamaterials: Determining parameters in the relaxed micromorphic model. *Proc. R. Soc. Lond. A* **472**, 20160169 (2016)
- J. Mandel, *Plasticité Classique et Viscoplasticité*, CISM Courses and Lectures No. 97, Udine (Springer, Berlin, 1971)
- J. Mandel, Equations constitutives et directeurs dans les milieux plastiques et viscoplastiques. *Int. J. Solids Struct.* **9**, 725–740 (1973)
- N. Marchal, S. Flouriou, S. Forest, L. Remy, Crack–tip stress–strain fields in single crystal nickel–base superalloys at high temperature under cyclic loading. *Comput. Mater. Sci.* **37**, 42–50 (2006a)
- N. Marchal, S. Forest, L. Rémy, S. Duvinage, Simulation of fatigue crack growth in single crystal superalloys using local approach to fracture, in *Local Approach to Fracture, 9th European Mechanics of Materials Conference, Euromech–Mecamat*, ed. by J. Besson, D. Moinereau, D. Steglich (Presses de l’Ecole des Mines de Paris, Moret–sur–Loing, 2006b), pp. 353–358
- G. Maugin, The method of virtual power in continuum mechanics: Application to coupled fields. *Acta Mech.* **35**, 1–70 (1980)
- G. Maugin, Internal variables and dissipative structures. *J. Non-Equilib. Thermo-dyn.* **15**, 173–192 (1990)
- G. Maugin, *Thermomechanics of Nonlinear Irreversible Behaviors* (World Scientific, Singapore/River Edge, 1999)
- C. Miehe, Variational gradient plasticity at finite strains. Part I: Mixed potentials for the evolution and update problems of gradient-extended dissipative solids. *Comput. Methods Appl. Mech. Eng.* **268**, 677–703 (2014)
- C. Miehe, N. Apel, M. Lambrecht, Anisotropic additive plasticity in the logarithmic strain space: Modular kinematic formulation and implementation based on incremental minimization principles for standard materials. *Comp. Methods Appl. Mech. Eng.* **191**, 5383–5425 (2002)

- C. Miehe, S. Teichtmeister, F. Aldakheel, Phase-field modelling of ductile fracture: a variational gradient-extended plasticity-damage theory and its micromorphic regularization. *Philos. Trans. R. Soc. Lond. A* **374**(2066) (2016)
- R. Mindlin, Micro-structure in linear elasticity. *Arch. Ration. Mech. Anal.* **16**, 51–78 (1964)
- R. Mindlin, Second gradient of strain and surface-tension in linear elasticity. *Int. J. Solids Struct.* **1**, 417–438 (1965)
- H. Mühlhaus, *Continuum Models for Materials with Microstructure* (Wiley, Chichester, 1995)
- P. Neff, I. Ghiba, A. Madeo, L. Placidi, G. Rosi, A unifying perspective: The relaxed linear micromorphic continuum. *Contin. Mech. Thermodyn.* **26**, 639–681 (2014)
- R. Peerlings, On the role of moving elastic-plastic boundaries in strain gradient plasticity. *Model. Simul. Mater. Sci. Eng.* **15**, S109–S120 (2007)
- R. Peerlings, M. Geers, R. de Borst, W. Brekelmans, A critical comparison of non-local and gradient-enhanced softening continua. *Int. J. Solids Struct.* **38**, 7723–7746 (2001)
- R. Peerlings, T. Massart, M. Geers, A thermodynamically motivated implicit gradient damage framework and its application to brick masonry cracking. *Comput. Methods Appl. Mech. Eng.* **193**, 3403–3417 (2004)
- K. Pham, H. Amor, J.J. Marigo, C. Maurini, Gradient damage models and their use to approximate brittle fracture. *Int. J. Damage Mech.* **20**, 618–652 (2011)
- L. Poh, R. Peerlings, M. Geers, S. Swaddiwudhipong, An implicit tensorial gradient plasticity model - formulation and comparison with a scalar gradient model. *Int. J. Solids Struct.* **48**, 2595–2604 (2011)
- V.D. Rancourt, B. Appolaire, S. Forest, K. Ammar, Homogenization of viscoplastic constitutive laws within a phase field approach. *J. Mech. Phys. Solids* **88**, 35–48 (2016)
- R. Regueiro, On finite strain micromorphic elastoplasticity. *Int. J. Solids Struct.* **47**, 786–800 (2010)
- G. Rosi, N. Auffray, Anisotropic and dispersive wave propagation within strain-gradient framework. *Wave Motion* **63**, 120–134 (2016)
- K. Saanouni, M. Hamed, Micromorphic approach for finite gradient-elastoplasticity fully coupled with ductile damage: Formulation and computational aspects. *Int. J. Solids Struct.* **50**, 2289–2309 (2013)
- C. Sansour, A theory of the elastic-viscoplastic Cosserat continuum. *Arch. Mech.* **50**, 577–597 (1998a)
- C. Sansour, A unified concept of elastic-viscoplastic Cosserat and micromorphic continua. *Journal de Physique IV* **8**(Pr8), 341–348 (1998b)
- C. Sansour, S. Skatulla, H. Zbib, A formulation for the micromorphic continuum at finite inelastic strains. *Int. J. Solids Struct.* **47**, 1546–1554 (2010)
- A.V. Shotov, J. Ihlemann, Analysis of some basic approaches to finite strain elasto-plasticity in view of reference change. *Int. J. Plast.* **63**, 183–197 (2014)
- F. Sidoroff, A. Dogui, Some issues about anisotropic elastic-plastic models at finite strain. *Int. J. Solids Struct.* **38**, 9569–9578 (2001)
- J.C. Simo, C. Miehe, Associative coupled thermoplasticity at finite strains: Formulation, numerical analysis and implementation. *Comput. Methods Appl. Mech. Eng.* **98**, 41–104 (1992)
- H. Steeb, S. Diebels, A thermodynamic-consistent model describing growth and remodeling phenomena. *Comput. Mater. Sci.* **28**, 597–607 (2003)
- S. Toll, The dissipation inequality in hypoplasticity. *Acta Mech.* **221**, 39–47 (2011)
- C. Truesdell, W. Noll, The non-linear field theories of mechanics, in *Handbuch der Physik*, ed. by S. Flügge, reedition (Springer, Berlin/Heidelberg, 1965)
- C. Truesdell, R. Toupin, The classical field theories, in *Handbuch der Physik*, ed. by S. Flügge, vol. 3 (Springer, Berlin, 1960), pp. 226–793
- P. Ván, A. Berezovski, C. Papenfuss, Thermodynamic approach to generalized continua. *Contin. Mech. Thermodyn.* **26**, 403–420 (2014)
- J. Vignollet, S. May, R.D. Borst, C. Verhoosel, Phase-field models for brittle and cohesive fracture. *Meccanica* **49**, 2587–2601 (2014)

-
- B. Wcislo, L. Pamin, K. Kowalczyk-Gajewska, Gradient-enhanced damage model for large deformations of elastic-plastic materials. *Arch. Mech.* **65**, 407–428 (2013)
 - S. Wulfinghoff, E. Bayerschen, T. Böhlke, Conceptual difficulties in plasticity including the gradient of one scalar plastic field variable. *PAMM. Proc. Appl. Math. Mech.* **14**, 317–318 (2014)
 - H. Xiao, O.T. Bruhns, A. Meyers, Existence and uniqueness of the integrable–exactly hypoelastic equation and its significance to finite elasticity. *Acta Mech.* **138**, 31–50 (1999)



Higher Order Thermo-mechanical Gradient Plasticity Model: Nonproportional Loading with Energetic and Dissipative Components

16

George Z. Voyiadjis and Yooseob Song

Contents

Introduction	548
Principle of Virtual Power	550
Thermodynamic Formulation with Higher Order Plastic Strain Gradients	553
The Energetic and Dissipative Components of the Thermodynamic Microstresses	554
Helmholtz Free Energy and Energetic Thermodynamic Microstresses	557
Dissipation Potential and Dissipative Thermodynamic Microstresses	561
Flow Rule	564
Thermo-mechanical Coupled Heat Equation	565
Finite Element Implementation of the Strain Gradient Plasticity Model	566
Experimental Validation of the Strain Gradient Plasticity Model	569
Stretch-Passivation Problem	574
Numerical Results	575
Conclusions	591
References	593

Abstract

In this chapter, two cases of thermodynamic-based higher order gradient plasticity theories are presented and applied to the stretch-surface passivation problem for investigating the material behavior under the nonproportional loading condition. This chapter incorporates the thermal and mechanical responses of microsystems. It addresses phenomena such as size and boundary effects and in particular microscale heat transfer in fast-transient processes. The stored energy of cold work is considered in the development of the recoverable

G. Z. Voyiadjis (✉) · Y. Song
Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge,
LA, USA
e-mail: voyiadjis@eng.lsu.edu; ysong17@lsu.edu

counterpart of the free energy. The main distinction between the two cases is the presence of the dissipative higher order microstress quantities $\mathcal{S}_{ijk}^{\text{dis}}$. Fleck et al. (Soc. A-Math. Phys. **470**:2170, 2014, ASME **82**:7, 2015) noted that $\mathcal{S}_{ijk}^{\text{dis}}$ always gives rise to the stress jump phenomenon, which causes a controversial dispute in the field of strain gradient plasticity theory with respect to whether it is physically acceptable or not, under the nonproportional loading condition. The finite element solution for the stretch-surface passivation problem is also presented by using the commercial finite element package ABAQUS/standard (*User's Manual (Version 6.12)*, Dassault Systemes Simulia Corp., Providence, 2012) via the user-subroutine UEL. The model is validated by comparing with three sets of small-scale experiments. The numerical simulation part, which is largely composed of four subparts, is followed. In the first part, the occurrence of the stress jump phenomenon under the stretch-surface passivation condition is introduced in conjunction with the aforementioned three experiments. The second part is carried out in order to clearly show the results to be contrary to each other from the two classes of strain gradient plasticity models. An extensive parametric study is presented in the third part in terms of the effects of the various material parameters on the stress-strain response for the two cases of strain gradient plasticity models, respectively. The evolution of the free energy and dissipation potentials are also investigated at elevated temperatures. In the last part, the two-dimensional simulation is given to examine the gradient and grain boundary effect, the mesh sensitivity of the two-dimensional model, and the stress jump phenomenon. Finally, some significant conclusions are presented.

Keywords

Higher order gradient plasticity · Energetic · Dissipative · Stress jump · Non-proportional loading

Introduction

It is well known that the classical continuum plasticity theory cannot capture the size effect of the microstructure during the course of plastic deformation. Aifantis (1984) incorporated a material length scale into the conventional continuum plasticity model to capture the size effect and proposed a modified flow rule by including the gradient term $\beta \nabla^2 \varepsilon^p$ into the conventional flow rule as follows:

$$\tau_{\text{eff}} = R(\varepsilon^p) - \beta \nabla^2 \varepsilon^p \quad (1)$$

where τ_{eff} is the effective stress and is calculated by $\tau_{\text{eff}} = \sqrt{3\tau_{ij}\tau_{ij}/2}$ with the deviatoric stress tensor τ_{ij} , $R(\varepsilon^p) > 0$ is the conventional flow resistance, ε^p is the accumulated plastic strain, $\beta > 0$ is a material coefficient, and $\nabla^2 = \text{Div}\nabla$ is the Laplacian operator. Hereafter, a number of mechanisms associated with geometrically necessary dislocations (GNDs) have been proposed within the framework of the strain gradient plasticity (SGP) theory. Generally, there are two different

kinds of viewpoints in SGP theory in terms of the origin of the strengthening effect. Firstly, there is a specific argument that the strain gradient strengthening is purely energetic in the sense that GNDs originate in the blockage by grain boundaries and the pile-ups of dislocations has a backstress associated with the energetic strengthening (Fleck and Willis 2015). For example, Kuroda and Tvergaard (2010) argued that the term $\tau_{\text{eff}} + \beta \nabla^2 \varepsilon^p$ calculated in Eq. 1 represents a total stress at the material point that activates the plastic straining, i.e., the generation and movement of dislocations. They also pointed out that the thermodynamic requirement on the plastic dissipation \mathcal{D} is evaluated by $\mathcal{D} = (\tau_{\text{eff}} + \beta \nabla^2 \varepsilon^p) \dot{\varepsilon}^p > 0$. This demonstration shows that the nonlocal term $\beta \nabla^2 \varepsilon^p$ is naturally interpreted as an energetic quantity, which is consistent with the interpretation in Gurtin and Anand (2009) that the nonlocal term in Aifantis' formulation should be energetic. Secondly, there is another point of view that GNDs combine with the statistically stored dislocations (SSDs) to provide the forest hardening, which in turn, lead to the dissipative strengthening. For example, in Fleck and Hutchinson (2001), gradient term is implicitly considered as a dissipative quantity that causes the theory to violate the thermodynamic requirement on plastic dissipation. Fleck and Willis (2009a) developed a mathematical basis for phenomenological gradient plasticity theory corresponding to both rate dependent/independent behavior with the scalar plastic multiplier. The plastic work in Fleck and Willis (2009a) is taken to be purely dissipative in nature, and the thermodynamic microstresses are assumed to be dissipative. In their incremental form of plasticity theory, an associated plastic flow rule is assumed by means of the convex yield function, consequently, the positive plastic work is ensured. Fleck and Willis (2009b) developed a phenomenological flow theory version of SGP theory by extending their theory in Fleck and Willis (2009a) to isotropic and anisotropic solids with tensorial plastic multiplier. Fleck and Willis (2009b) argued that the microstress quantities should include a dissipative part; thus, it has been proposed that the term $\beta \nabla^2 \varepsilon^p$ is additively decomposed into an energetic $(\cdot)_{\text{en}}$ and dissipative $(\cdot)_{\text{dis}}$ in order to develop a kinematic hardening theory. The dissipative stresses satisfy a yield condition with an associated flow plastic rule, while the free energy provides the standard kinematic hardening.

There has been a debate between Fleck, Willis, and Hutchinson (Fleck et al. 2014, 2015; Hutchinson 2012) and Gudmundson et al. (Gudmundson 2004; Gurtin and Anand 2009) for the last 15 years or so. Fleck and Hutchinson (2001) developed a phenomenological SGP theory using higher order tensors with a similar framework to that proposed by Aifantis (1984) and Muhlhaus and Aifantis (1991). Higher order stresses and additional boundary conditions have been involved in the theory to develop a generalization of the classical rate-independent J_2 flow theory of gradient plasticity. However, they do not discuss the compatibility of their theory with thermodynamic requirements on the plastic dissipation. Gudmundson (2004) and Gurtin and Anand (2009) pointed out that the formulation of Fleck and Hutchinson (2001) violates thermodynamic requirements on the plastic dissipation. Gurtin and Anand (2009) discussed the physical nature of nonlocal terms in the flow rules developed by Fleck and Hutchinson (2001) under isothermal condition and concluded that the flow rule of Fleck and Hutchinson (2001) does not always satisfy

the thermodynamic requirement on plastic dissipation unless the nonlocal term is dropped. A formulation of Fleck and Hutchinson (2001) has been modified to meet this thermodynamic requirement by partitioning the higher order microstresses into energetic and dissipative components (Hutchinson 2012). In addition, Hutchinson (2012) classified the strain gradient version of J_2 flow theories into two classes: incremental theory developed by Fleck and Hutchinson and nonincremental theory developed by Gudmundson et al. (c.f. see Fleck et al. (2014, 2015), Gudmundson (2004), Gurtin and Anand (2005, 2009), and Hutchinson (2012) for details). The specific phenomenon in the nonincremental theory that exhibits a finite stress jump due to infinitesimal changes in plastic strain that may occur under the nonproportional loading is noted and its physical acceptance is also discussed in the work of Fleck et al. (2014, 2015). Hutchinson (2012) concluded that discontinuous changes with infinitesimal changes in boundary loads are physically suspect. (Despite the argument of J.W. Hutchinson, *Acta Mech. Sin.* 28, 4 (2012), there is another viewpoint to look at the stress jump phenomenon. The perspective taken in N.A. Fleck, J.R. Willis, *ibid.* 31, (2015) is that it is premature at this moment in time to judge whether a formulation associated with the stress jump is physically acceptable or not, therefore, an in-depth study of dislocation mechanism and microscale experiments with non-proportional loading history is needed.) Fleck et al. (2014, 2015) have shown this phenomenon with two plane strain problems, stretch-passivation problem, and stretch-bending problem, for nonproportional loading condition. In their work, it is noted that the dissipative higher order microstress quantities $\mathcal{S}_{ijk}^{\text{dis}}$ always generate the stress jump for nonproportional loading problems.

In this chapter, two different cases of the high order SGP model with and without the dissipative higher order microstress quantities $\mathcal{S}_{ijk}^{\text{dis}}$ are presented based on the new forms of the free energy and the dissipation potentials for eliminating an elastic loading gap. The presented model is applied to the stretch-surface passivation problem in order to compare the behavior of each case under the nonproportional loading condition. For this, the finite element solution for the stretch-surface passivation problem is presented by using the commercial finite element package ABAQUS/standard (2012) via the user-subroutine UEL and validated by comparing with three sets of small-scale experiments, which have been conducted by Han et al. (2008), Haque and Saif (2003), and Xiang and Vlassak (2006). An extensive numerical work is also carried out based on the one-dimensional and two-dimensional codes in order to compare the results from the two cases of the SGP model and to analyze the characteristics of the stress jump phenomenon.

Principle of Virtual Power

The principle of virtual power is used to derive the local equation of motion and the nonlocal microforce balance for volume V as well as the equations for local traction force and nonlocal microtraction condition for the external surface S , respectively. In the presence of varying temperature fields at the microstructure level, the formulation should incorporate the effects of the temperature gradient

on the thermo-mechanical behavior of the material due to the microheterogeneous nature of material (Forest and Amestoy 2008). In this sense, it is assumed here that the plastic strain, plastic strain gradient, temperature, and temperature gradient contribute to the power per unit volume.

Moreover, as it is mentioned in section “Introduction,” the effect of the interface plays a crucial role for the plastic behavior of the material at the microscale. An interface (grain boundary) separating grains \mathcal{G}_1 and \mathcal{G}_2 is taken into account here, and it is assumed that the displacement field is continuous, i.e. $u_i^{\mathcal{G}_1} = u_i^{\mathcal{G}_2}$, across the grain boundary (Fig. 1). As shown in this figure, a dislocation moving toward the grain boundary in grain \mathcal{G}_1 cannot pass through the grain boundary, but it is trapped and accumulated at the grain boundary due to the misalignment of the grains \mathcal{G}_1 and \mathcal{G}_2 that are contiguous to each other. In this sense, the grain boundary acts as an obstacle to block the dislocation movement; therefore, the yield strength of the material increases as the number of grain boundaries increases. By assuming that the interface surface energy depends on the plastic strain rate at the interface of the plastically deforming phase, the internal part of the principle of virtual power for the bulk P_{int} and for the interface P_{int}^I are expressed in terms of the energy contributions in the arbitrary subregion of the volume V and the arbitrary subsurface of the interface S^I , respectively, as follows:

$$P_{\text{int}} = \int_V \left(\sigma_{ij} \dot{\varepsilon}_{ij}^e + \chi_{ij} \dot{\varepsilon}_{ij}^p + \mathcal{S}_{ijk} \dot{\varepsilon}_{ij,k}^p + \mathcal{A}\dot{T} + \mathcal{B}_i \dot{T}_{,i} \right) dV \quad (2)$$

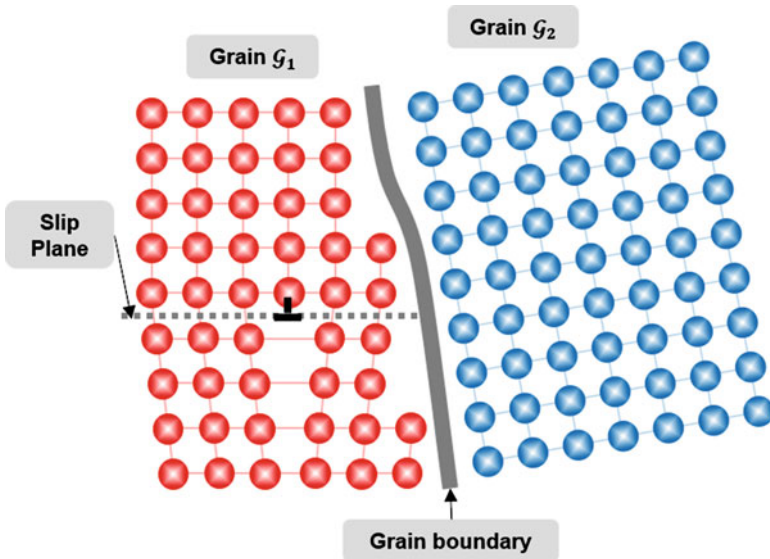


Fig. 1 The schematic illustration of the spatial lattice of two contiguous grains, \mathcal{G}_1 and \mathcal{G}_2 , along with a single slip in grain \mathcal{G}_1 (Reprinted with permission from Voyiadjis et al. 2017)

$$P_{\text{int}}^I = \int_{S^I} \left(\mathbb{M}_{ij}^{IG_1} \dot{\varepsilon}_{ij}^{pIG_1} + \mathbb{M}_{ij}^{IG_2} \dot{\varepsilon}_{ij}^{pIG_2} \right) dS^I \quad (3)$$

where the superscripts “e,” p ,” and “ I ” are used to express the elastic state, the plastic state, and the interface, respectively. The internal power for the bulk in the form of Eq. 2 is defined using the Cauchy stress tensor σ_{ij} , the backstress \mathcal{X}_{ij} conjugate to the plastic strain rate $\dot{\varepsilon}_{ij}^p$, the higher order microstress \mathcal{S}_{ijk} conjugate to the gradients of the plastic strain rate $\dot{\varepsilon}_{ij,k}^p$, and the generalized stresses \mathcal{A} and \mathcal{B}_i conjugate to the temperature rate \dot{T} and the gradient of the temperature rate $\dot{T}_{,i}$, respectively. The internal power for the interface in the form of Eq. 3 is defined using the interfacial microscopic moment tractions $\mathbb{M}_{ij}^{IG_1}$ and $\mathbb{M}_{ij}^{IG_2}$ expending power over the interfacial plastic strain rates $\dot{\varepsilon}_{ij}^{pIG_1}$ at S^{IG_1} and $\dot{\varepsilon}_{ij}^{pIG_2}$ at S^{IG_2} , respectively.

Moreover, since the plastic deformation, which is accommodated by the generation and motion of the dislocation, is influenced by the interfaces, Eq. 3 results in higher order boundary conditions generally consistent with the framework of a gradient type theory. These extra boundary conditions should be imposed at internal and external boundary surfaces or interfaces between neighboring grains (Aifantis and Willis 2005; Gurtin 2008). The internal power for the bulk is balanced with the external power for the bulk expended by the tractions on the external surfaces S and the body forces acting within the volume V as shown below:

$$P_{\text{ext}} = \int_V b_i v_i dV + \int_S \left(t_i v_i + m_{ij} \dot{\varepsilon}_{ij}^p + a \dot{T} \right) dS \quad (4)$$

where t_i and b_i are traction and the external body force conjugate to the macroscopic velocity v_i , respectively. It is further assumed here that the external power has terms with the microtraction m_{ij} and a , conjugate to the plastic strain rate $\dot{\varepsilon}_{ij}^p$ and the temperature rate \dot{T} , respectively, since the internal power in Eq. 2 contains the terms of the gradients of the plastic strain rate $\dot{\varepsilon}_{ij,k}^p$ and the gradients of the temperature rate $\dot{T}_{,i}$, respectively.

Making use of the principle of virtual power that the external power is equal to the internal power ($P_{\text{int}} = P_{\text{ext}}$) along with the integration by parts on some terms in Eq. 2, and utilizing the divergence theorem, the equations for balance of linear momentum and nonlocal microforce balance can be represented, respectively, for volume V , as follows:

$$\sigma_{ij,j} + b_i = 0 \quad (5)$$

$$\mathcal{X}_{ij} - \tau_{ij} - \mathcal{S}_{ijk,k} = 0 \quad (6)$$

$$\text{div } \mathcal{B}_i - \mathcal{A} = 0 \quad (7)$$

where $\tau_{ij} = \sigma_{ij} - \sigma_{kk} \delta_{ij} / 3$ is the deviatoric part of the Cauchy stress tensor and δ_{ij} is the Kronecker delta.

On the external surface S , the equations for local surface traction conditions and nonlocal microtraction conditions can be written, respectively, with the outward unit normal vector to S , n_k , as follows:

$$t_j = \sigma_{ij} n_i \quad (8)$$

$$m_{ij} = \mathcal{S}_{ijk} n_k \quad (9)$$

$$a = \mathcal{B}_i n_i \quad (10)$$

In addition, the interfacial external power P_{ext}^I , which is balanced with the interfacial internal power P_{int}^I , is expended by the macrotractions $\sigma_{ij}^{\mathcal{G}_1}(-n_j^I)$ and $\sigma_{ij}^{\mathcal{G}_2}(n_j^I)$ conjugate to the macroscopic velocity v_i , and the microtractions $\mathbb{S}_{ijk}^{I\mathcal{G}_1}(-n_k^I)$ and $\mathbb{S}_{ijk}^{I\mathcal{G}_2}(n_k^I)$ that are conjugate to $\dot{\varepsilon}_{ij}^{I\mathcal{G}_1}$ and $\dot{\varepsilon}_{ij}^{I\mathcal{G}_2}$, respectively, as follows:

$$P_{\text{ext}}^I = \int_{S^I} \left\{ \left(\sigma_{ij}^{\mathcal{G}_2} n_j^I - \sigma_{ij}^{\mathcal{G}_1} n_j^I \right) v_i + \mathbb{S}_{ijk}^{I\mathcal{G}_2} n_k^I \dot{\varepsilon}_{ij}^{I\mathcal{G}_2} - \mathbb{S}_{ijk}^{I\mathcal{G}_1} n_k^I \dot{\varepsilon}_{ij}^{I\mathcal{G}_1} \right\} dS^I \quad (11)$$

By equating $P_{\text{int}}^I = P_{\text{ext}}^I$ with considering the arbitrary variation of the plastic strain at the interface, the interfacial macro- and microforce balances can be obtained as follows:

$$\left(\sigma_{ij}^{\mathcal{G}_1} - \sigma_{ij}^{\mathcal{G}_2} \right) n_j^I = 0 \quad (12)$$

$$\mathbb{M}_{ij}^{I\mathcal{G}_1} + \mathbb{S}_{ijk}^{I\mathcal{G}_1} n_k^I = 0 \quad (13)$$

$$\mathbb{M}_{ij}^{I\mathcal{G}_2} - \mathbb{S}_{ijk}^{I\mathcal{G}_2} n_k^I = 0 \quad (14)$$

The microforce balance conditions in Eqs. 13 and 14 represent the coupling behavior in the grain interior at the interface to the behavior of the interface, since the microtractions $\mathbb{S}_{ijk}^{I\mathcal{G}_1} n_k^I$ and $\mathbb{S}_{ijk}^{I\mathcal{G}_2} n_k^I$ are the special cases of Eq. 8 for the internal surface of the interface.

Thermodynamic Formulation with Higher Order Plastic Strain Gradients

The first law of thermodynamics, which encompasses several principles including the law of conservation of energy, is taken into account in this chapter in order to develop a thermodynamically consistent formulation accounting for the thermo-mechanical behavior of small-scale metallic volumes during the fast transient process. In order to consider micromechanical evolution in the first law of thermodynamics, the enhanced SGP theory with the plastic strain gradient is employed for the

mechanical part of the formulation, whereas the micromorphic model is employed for the thermal counterpart as follows (see the work of Forest and co-workers (Forest and Amestoy 2008; Forest and Sievert 2003)):

$$\rho \dot{e} = \sigma_{ij} \dot{\epsilon}_{ij}^e + \mathcal{X}_{ij} \dot{\epsilon}_{ij}^p + \mathcal{S}_{ijk} \dot{\epsilon}_{ij,k}^p + \mathcal{A} \dot{T} + \mathcal{B}_i \dot{T}_{,i} - \text{div} q_i \quad (15)$$

$$\dot{e}^I = \mathbb{M}_{ij}^I \dot{\epsilon}_{ij}^{pI} + q_i^I n_i^I \quad (16)$$

where ρ is the mass density, e is the specific internal energy, e^I is the internal surface energy density at the contacting surface, and q_i and q_i^I are the heat flux vectors of the bulk and the interface, respectively.

The second law of thermodynamics, or entropy production inequality as it is often called, yields a physical basis that accounts for the distribution of GNDs within the body along with the energy carrier scattering and requires that the free energy increases at a rate not greater than the rate at which work is performed. Based on this requirement, entropy production inequalities for the bulk and the interface can be expressed, respectively, as follows:

$$\rho \dot{s} T - \rho \dot{e} + \sigma_{ij} \dot{\epsilon}_{ij}^e + \mathcal{X}_{ij} \dot{\epsilon}_{ij}^p + \mathcal{S}_{ijk} \dot{\epsilon}_{ij,k}^p - q_i \frac{T_{,i}}{T} + \mathcal{A} \dot{T} + \mathcal{B}_i \dot{T}_{,i} \geq 0 \quad (17)$$

$$\dot{s}^I T^I - q_i^I n_i^I \geq 0 \quad (18)$$

where s is the specific entropy and s^I is the surface density of the entropy of the interface.

The Energetic and Dissipative Components of the Thermodynamic Microstresses

Internal energy e , temperature T , and entropy s describing the current state of the material can be attributed to the Helmholtz free energy Ψ (per unit volume) such as

$$\rho \dot{s} T - \rho \dot{e} + \sigma_{ij} \dot{\epsilon}_{ij}^e + \mathcal{X}_{ij} \dot{\epsilon}_{ij}^p + \mathcal{S}_{ijk} \dot{\epsilon}_{ij,k}^p - q_i \frac{T_{,i}}{T} + \mathcal{A} \dot{T} + \mathcal{B}_i \dot{T}_{,i} \geq 0. \quad (19)$$

By taking the time derivative of Eq. 19 for the bulk and the interface and substituting each into Eqs. 17 and 18, respectively, the nonlocal free energy (i.e., Clausius-Duhem) inequality for the bulk and the interface can be obtained as follows:

$$\sigma_{ij} \dot{\epsilon}_{ij}^e + \mathcal{X}_{ij} \dot{\epsilon}_{ij}^p + \mathcal{S}_{ijk} \dot{\epsilon}_{ij,k}^p + \mathcal{A} \dot{T} + \mathcal{B}_i \dot{T}_{,i} - \rho \dot{\Psi} - \rho \dot{s} T - q_i \frac{T_{,i}}{T} \geq 0 \quad (20)$$

$$\mathbb{M}_{ij}^I \dot{\epsilon}_{ij}^{pI} - \dot{\Psi}^I - s^I \dot{T}^I \geq 0 \quad (21)$$

In order to derive the constitutive equations within a small-scale framework, an attempt to account for the effect of nonuniform distribution of the microdefection with temperature on the homogenized response of the material is carried out in this chapter with the functional forms of the Helmholtz free energy in terms of its state variables. By assuming the isothermal condition for the interface (i.e., $\dot{T}^I = 0$), the Helmholtz free energy for the bulk and the interface are given, respectively, by

$$\Psi = \Psi \left(\varepsilon_{ij}^e, \varepsilon_{ij}^p, \varepsilon_{ij,k}^p, T, T_i \right) \quad (22)$$

$$\Psi^I = \Psi^I \left(\varepsilon_{ij}^{pI} \right) \quad (23)$$

where the function Ψ is assumed to be smooth and the function Ψ^I is assumed to be convex with respect to a plastic strain at the interface ε_{ij}^{pI} .

Taking time derivative of the Helmholtz free energy for the bulk $\dot{\Psi}$ and the interface $\dot{\Psi}^I$ give the following expressions, respectively

$$\dot{\Psi} = \frac{\partial \Psi}{\partial \varepsilon_{ij}^e} \dot{\varepsilon}_{ij}^e + \frac{\partial \Psi}{\partial \varepsilon_{ij}^p} \dot{\varepsilon}_{ij}^p + \frac{\partial \Psi}{\partial \varepsilon_{ij,k}^p} \dot{\varepsilon}_{ij,k}^p + \frac{\partial \Psi}{\partial T} \dot{T} + \frac{\partial \Psi}{\partial T_i} \dot{T}_i \quad (24)$$

$$\dot{\Psi}^I = \frac{\partial \Psi^I}{\partial \varepsilon_{ij}^{pI}} \dot{\varepsilon}_{ij}^{pI} \quad (25)$$

By substituting Eq. 24 into Eq. 20 for the bulk and Eq. 25 into Eq. 21 for the interface, and factoring out the common terms, one obtains the following inequalities:

$$\begin{aligned} & \left(\sigma_{ij} - \rho \frac{\partial \Psi}{\partial \varepsilon_{ij}^e} \right) \dot{\varepsilon}_{ij}^e + \left(\mathcal{X}_{ij} - \rho \frac{\partial \Psi}{\partial \varepsilon_{ij}^p} \right) \dot{\varepsilon}_{ij}^p + \left(\mathcal{S}_{ijk} - \rho \frac{\partial \Psi}{\partial \varepsilon_{ij,k}^p} \right) \dot{\varepsilon}_{ij,k}^p \\ & + \left(\mathcal{A} - \rho s - \rho \frac{\partial \Psi}{\partial T} \right) \dot{T} + \left(\mathcal{B}_i - \rho \frac{\partial \Psi}{\partial T_i} \right) \dot{T}_i - \frac{q_i}{T} T_i \geq 0 \end{aligned} \quad (26)$$

$$\mathbb{M}_{ij}^I \dot{\varepsilon}_{ij}^{pI} - \rho \frac{\partial \Psi^I}{\partial \varepsilon_{ij}^{pI}} \dot{\varepsilon}_{ij}^{pI} \geq 0 \quad (27)$$

Guided by Eqs. 26 and 27, it is further assumed that the thermodynamic microstress quantities \mathcal{X}_{ij} , \mathcal{S}_{ijk} , \mathcal{A} , and \mathbb{M}_{ij}^I are decomposed into the energetic and the dissipative components such as (Voyiadjis and Deliktas 2009; Voyiadjis and Faghihi 2012; Voyiadjis et al. 2014):

$$\mathcal{X}_{ij} = \mathcal{X}_{ij}^{\text{en}} + \mathcal{X}_{ij}^{\text{dis}} \quad (28)$$

$$\mathcal{S}_{ijk} = \mathcal{S}_{ijk}^{\text{en}} + \mathcal{S}_{ijk}^{\text{dis}} \quad (29)$$

$$\mathcal{A} = \mathcal{A}^{\text{en}} + \mathcal{A}^{\text{dis}} \quad (30)$$

$$\mathbb{M}_{ij}^I = \mathbb{M}_{ij}^{I,\text{en}} + \mathbb{M}_{ij}^{I,\text{dis}} \quad (31)$$

The components $\mathbb{M}_{ij}^{I,\text{en}}$ and $\mathbb{M}_{ij}^{I,\text{dis}}$ represent the mechanisms for the pre- and postslip transfer and thus involve the plastic strain at the interface prior to the slip transfer $\varepsilon_{ij}^{pI(\text{pre})}$ and the one after the slip transfer $\varepsilon_{ij}^{pI(\text{post})}$, respectively. The overall plastic strain at the interface can be obtained by the summation of both plastic strains such as:

$$\varepsilon_{ij}^{pI} = \varepsilon_{ij}^{pI(\text{pre})} + \varepsilon_{ij}^{pI(\text{post})} \quad (32)$$

Substituting Eqs. 28, 29, and 30 into Eq. 26 for the bulk and Eq. 31 into Eq. 27 for the interface and rearranging them in accordance with the energetic and the dissipative parts give the following expressions:

$$\begin{aligned} & \left(\sigma_{ij} - \rho \frac{\partial \Psi}{\partial \varepsilon_{ij}^e} \right) \dot{\varepsilon}_{ij}^e + \left(\mathcal{X}_{ij}^{\text{en}} - \rho \frac{\partial \Psi}{\partial \varepsilon_{ij}^p} \right) \dot{\varepsilon}_{ij}^p + \left(\mathcal{S}_{ijk}^{\text{en}} - \rho \frac{\partial \Psi}{\partial \varepsilon_{ijk}^p} \right) \dot{\varepsilon}_{ij,k}^p \\ & + \left(\mathcal{A}^{\text{en}} - \rho s - \rho \frac{\partial \Psi}{\partial T} \right) \dot{T} + \left(\mathcal{B}_i - \rho \frac{\partial \Psi}{\partial T_i} \right) \dot{T}_i + \mathcal{X}_{ij}^{\text{dis}} \dot{\varepsilon}_{ij}^p \\ & + \mathcal{S}_{ijk}^{\text{dis}} \dot{\varepsilon}_{ij,k}^p + \mathcal{A}^{\text{dis}} \dot{T} - \frac{q_i}{T} T_i \geq 0 \end{aligned} \quad (33)$$

$$\left(\mathbb{M}_{ij}^{I,\text{en}} - \rho \frac{\partial \Psi^I}{\partial \varepsilon_{ij}^{pI}} \right) \dot{\varepsilon}_{ij}^{pI} + \mathbb{M}_{ij}^{I,\text{dis}} \dot{\varepsilon}_{ij}^{pI} \geq 0 \quad (34)$$

From the above equations with the assumption that the fifth term in Eq. 33 is strictly energetic, one can retrieve the definition of the energetic part of the thermodynamic microstresses as follows:

$$\begin{aligned} \sigma_{ij} &= \rho \frac{\partial \Psi}{\partial \varepsilon_{ij}^e}; \mathcal{X}_{ij}^{\text{en}} = \rho \frac{\partial \Psi}{\partial \varepsilon_{ij}^p}; \mathcal{S}_{ijk}^{\text{en}} = \rho \frac{\partial \Psi}{\partial \varepsilon_{ijk}^p}; \\ \mathcal{A}^{\text{en}} &= \rho \left(s + \frac{\partial \Psi}{\partial T} \right); \mathcal{B}_i = \rho \frac{\partial \Psi}{\partial T_i} \end{aligned} \quad (35)$$

$$\mathbb{M}_{ij}^{I,\text{en}} = \rho \frac{\partial \Psi^I}{\partial \varepsilon_{ij}^{pI}} \quad (36)$$

Hence, the residual respective dissipation is then obtained as:

$$\mathcal{D} = \mathcal{X}_{ij}^{\text{dis}} \dot{\varepsilon}_{ij}^p + \mathcal{S}_{ijk}^{\text{dis}} \dot{\varepsilon}_{ij,k}^p + \mathcal{A}^{\text{dis}} \dot{T} - \frac{q_i}{T} T_i \geq 0 \quad (37)$$

$$\mathcal{D}^I = \mathbb{M}_{ij}^{I,\text{dis}} \dot{\varepsilon}_{ij}^{pI} \geq 0 \quad (38)$$

where \mathcal{D} and \mathcal{D}^I are the dissipation densities per unit time for the bulk and the interface, respectively. The definition of the dissipative thermodynamic

microstresses can be obtained from the complementary part of the dissipation potentials $\mathcal{D}(\dot{\varepsilon}_{ij}^p, \dot{\varepsilon}_{ij,k}^p, \dot{T}, T_{,i})$ and $\mathcal{D}^I(\dot{\varepsilon}_{ij}^{pI})$ such as

$$\chi_{ij}^{\text{dis}} = \frac{\partial \mathcal{D}}{\partial \dot{\varepsilon}_{ij}^p}; \mathcal{S}_{ijk}^{\text{dis}} = \frac{\partial \mathcal{D}}{\partial \dot{\varepsilon}_{ij,k}^p}; \mathcal{A}^{\text{dis}} = \frac{\partial \mathcal{D}}{\partial \dot{T}}; -\frac{q_i}{T} = \frac{\partial \mathcal{D}}{\partial T_{,i}} \quad (39)$$

$$\mathbb{M}_{ij}^{I,\text{dis}} = \frac{\partial \mathcal{D}^I}{\partial \dot{\varepsilon}_{ij}^{pI}} \quad (40)$$

One now proceeds to present the constitutive laws for both the energetic and the dissipative parts which are achieved by employing the free energy and the dissipation potentials, which relate the stresses to their work-conjugate generalized stresses. The functional forms of the Helmholtz free energy and dissipation potential and the corresponding energetic and dissipative thermodynamic microstresses for the aforementioned two different cases of the model, i.e., the case with the dissipative higher order microstress quantities $\mathcal{S}_{ijk}^{\text{dis}}$ and the one without $\mathcal{S}_{ijk}^{\text{dis}}$, are presented in the following sections.

Helmholtz Free Energy and Energetic Thermodynamic Microstresses

Defining a specific form of the Helmholtz free energy function Ψ is tremendously important since it constitutes the bases in deriving the constitutive equations. In this chapter, the Helmholtz free energy function is put forward with three main counterparts, i.e., elastic, defect, and thermal energy, as follows (Voyiadjis and Song 2017):

$$\Psi = \underbrace{\left\{ \frac{1}{2\rho} \varepsilon_{ij}^e E_{ijkl} \varepsilon_{kl}^e - \frac{\alpha_t}{\rho} (T - T_r) \varepsilon_{ij}^e \delta_{ij} \right\}}_{\Psi^e(\text{elastic})} + \underbrace{\left\{ \frac{1}{\rho} \left(\frac{h}{r+1} \left(1 - \left(\frac{T}{T_y} \right)^n \right) \varepsilon_p^{r+1} + \frac{G}{2} \ell_{en}^2 \varepsilon_{lm,n}^p \varepsilon_{lm,n}^p \right) \right\}}_{\Psi^d(\text{defect})} + \underbrace{\left\{ -\frac{1}{2} \frac{c_\varepsilon}{T_r} (T - T_r)^2 - \frac{1}{2\rho} a T_{,i} T_{,i} \right\}}_{\Psi^t(\text{thermal})} \quad (41)$$

where E_{ijkl} is the elastic modulus tensor, α_t is the coefficient of linear thermal expansion, T_r is the reference temperature, h is the hardening material constant

corresponding to linear kinematic hardening, r ($0 < r < 1$) is the isotropic hardening material constant, $\varepsilon_p = \sqrt{\varepsilon_{ij}^p \varepsilon_{ij}^p}$ is the accumulated plastic strain, T_y and n are the thermal material constants that need to be calibrated by comparing with the experimental data, ℓ_{en} is the energetic length scale that describes the feature of the short-range interaction of GNDs, G is the shear modulus for isotropic linear elasticity, c_e is the specific heat capacity at constant stress, and a is a material constant accounting for the interaction between energy carriers such as phonon-electron. The term $(1 - (T/T_y)^n)$ in Eq. 41 accounts for the thermal activation mechanism for overcoming the local obstacles to dislocation motion.

The first term of the defect energy Ψ_1^d characterizes the interaction between slip systems, i.e., the forest dislocations leading to isotropic hardening. This term is further assumed to be decomposed into the recoverable counterpart $\Psi_1^{d,R}$ and nonrecoverable counterpart $\Psi_1^{d,NR}$. The establishment of the plastic strain gradient independent stored energy of cold work with no additional material parameters is achievable with this decomposition.

The recoverable counterpart, $\Psi_1^{d,R}$, accounts for the stored energy of cold work. When the elasto-plastic solid is cold-worked, most of the mechanical energy is converted into heat, but the remaining contributes to the stored energy of cold work through the creation and rearrangement of crystal defects such as dislocations, point defects, line defects, and stacking faults (Rosakis et al. 2000). In this chapter, the plastic strain-dependent free energy, $\Psi_1^{d,R}$, accounting for the stored energy of cold work is derived by assuming that the stored energy is related to the energy carried by dislocations. Mollica et al. (2001) investigated the inelastic behavior of the metals subject to loading reversal by linking the hardening behavior of the material to thermo-dynamical quantities such as the stored energy due to cold work and the rate of dissipation. In the work of Mollica et al. (2001), it is assumed that the stored energy depends on the density of the dislocation network that increases with monotonic plastic deformation until it is saturated at some point. This points out that the material stores this energy for a certain range of the accumulated plastic strain, after which the material will mainly dissipate the external work supply.

For the derivation of the stored energy of cold work, one assumes that the energetic microstress χ_{ij}^{en} , given later by Eq. 52, can be expressed by separation of variables as follows:

$$\chi_{ij}^{en}(\varepsilon_{ij}^p, T) = \sum(\varepsilon_{ij}^p) \mathcal{T}(T) \quad (42)$$

with $\sum(\varepsilon_{ij}^p) = h \varepsilon_{ij}^p \varepsilon_p^{r-1}$ and $\mathcal{T}(T) = (1 - (T/T_y)^n)$.

On the other hand, instead of using the plastic strain at the macroscale level to describe the plastic deformation, Σ can be defined at the microscale level using the Taylor law, which gives a simple relation between the shear strength and the dislocation density, as follows:

$$\Sigma = \varsigma G b \sqrt{\rho_t} \quad (43)$$

where ζ is the statistical coefficient accounting for the deviation from regular spatial arrangements of the dislocation populations, b is the magnitude of the Burgers vector, and ρ_t is the equivalent total dislocation density and can be obtained by $(\Sigma/\zeta Gb)^2$ from Eq. 43.

Here, it is assumed that the stored energy of cold work, the result of energy carried by each dislocation, results in an extra latent hardening which is recoverable and temperature independent. Thus, the recoverable energy of cold work can be put forward as follows:

$$\Psi_1^{d,R} = \mathbb{U} \rho_t \quad (44)$$

where \mathbb{U} is the elastic deformation energy of a dislocation and can be approximately given by

$$\mathbb{U} = \frac{Gb^2}{4\pi} \ln \left(\frac{R}{R_0} \right) \quad (45)$$

where R is the cut-off radius ($R \approx 10^3 b$) and R_0 is the internal radius ($b < R_0 < 10b$) (Meyers and Chawla 2009). By substituting $\rho_t = (\Sigma/\zeta Gb)^2$ into Eq. 45 along with Eq. 42, one can obtain the stored energy of cold work as follows:

$$\Psi_1^{d,R} = \vartheta h^2 \varepsilon_p^{2r} \quad (46)$$

where by comparing the aforementioned ranges for R , R_0 , and ζ to the shear modulus G , ϑ can be expressed by

$$\vartheta = \frac{1}{4\pi\zeta^2 G} \ln \left(\frac{R}{R_0} \right) \approx \frac{1}{G} \quad (47)$$

The nonrecoverable counterpart $\Psi_1^{d,NR}$ accounting for the energetically based hardening rule that mimics the dissipative behavior by describing irreversible loading processes can then be derived as follows (Gurtin and Reddy 2009):

$$\Psi_1^{d,NR} = \frac{h}{r+1} \left(1 - \left(\frac{T}{T_y} \right)^n \right) \varepsilon_p^{r+1} - \vartheta h^2 \varepsilon_p^{2r} \quad (48)$$

where ϑ is a constant that depends on the material microstructure.

The second term of the defect energy Ψ_2^d characterizes the short-range interactions between coupling dislocations moving on close slip planes and leads to the kinematic hardening. This defect energy Ψ_2^d is recoverable in the sense that by starting at any value of the accumulated plastic strain gradients, Ψ_2^d returns to its original value as the accumulated plastic strain gradients return to their original value.

One can now obtain the energetic thermodynamic forces by using the definitions in Eq. 35 along with the Helmholtz free energy given by Eq. 41 as follows:

$$\sigma_{ij} = E_{ijkl}\varepsilon_{kl}^e - \alpha_t (T - T_r) \delta_{ij} \quad (49)$$

$$\mathcal{A}^{\text{en}} = \rho s - \alpha_t (T - T_r) \varepsilon_{ij}^e \delta_{ij} - \frac{c_\varepsilon}{T_r} (T - T_r) - \frac{h\varepsilon_p^{r+1}}{r+1} \frac{n}{T_y} \left(\frac{T}{T_y}\right)^{n-1} \quad (50)$$

$$\mathcal{B}_i = -aT_{,i} \quad (51)$$

$$\mathcal{X}_{ij}^{\text{en}} = h\varepsilon_{ij}^p \left(1 - \left(\frac{T}{T_y}\right)^n\right) \varepsilon_p^{r-1} \quad (52)$$

$$\mathcal{S}_{ijk}^{\text{en}} = G\ell_{\text{cn}}^2 \varepsilon_{ij,k}^p \quad (53)$$

Here, according to the aforementioned decomposition of the first term of the defect energy into the recoverable and nonrecoverable counterparts, $\mathcal{X}_{ij}^{\text{en}}$ can be further decomposed into recoverable ($\mathcal{X}_{ij}^{\text{en},R}$) and nonrecoverable ($\mathcal{X}_{ij}^{\text{en},\text{NR}}$) counterparts as follows:

$$\mathcal{X}_{ij}^{\text{en},R} = 2r\vartheta h^2 \varepsilon_{ij}^p \varepsilon_p^{2r-2} \quad (54)$$

$$\mathcal{X}_{ij}^{\text{en},\text{NR}} = h\varepsilon_{ij}^p \left(1 - \left(\frac{T}{T_y}\right)^n\right) \varepsilon_p^{r-1} - 2r\vartheta h^2 \varepsilon_{ij}^p \varepsilon_p^{2r-2} \quad (55)$$

From the aforementioned physical interpretations of $\Psi_1^{d,R}$ and $\Psi_1^{d,\text{NR}}$, as well as $\mathcal{X}_{ij}^{\text{en},R}$ and $\mathcal{X}_{ij}^{\text{en},\text{NR}}$ can be defined as the terms describing the reversible loading due to the energy carried by dislocations and the energetically based hardening rule that mimics dissipative behavior by describing irreversible loading processes, respectively.

Meanwhile, it is well known that the interface plays a role as the barrier to plastic slip in the early stages of plastically deforming phase, while it acts as a source of the dislocation nucleation in the later stages. The energetic condition in the area around the interface is affected by the long-range internal stress fields associated with constrained plastic flow which leads to the accumulated and pile-up of dislocations near the interface. Thus, the condition at the interface is determined by a surface energy that depends on the plastic strain state at the interface (Fredriksson and Gudmundson 2005).

The interfacial Helmholtz free energy per unit surface area of the interface is put forward under the guidance of Fredriksson and Gudmundson (2005) work such as (It should be noted that it is possible to introduce another form of the surface energy if it is convex in, ε_{ij}^p):

$$\Psi^I = \frac{1}{2} G \ell_{\text{en}}^I \varepsilon_{ij}^{pI(\text{pre})} \varepsilon_{ij}^{pI(\text{pre})} \quad (56)$$

where ℓ_{en}^I is the interfacial recoverable length scale.

By substituting the interfacial Helmholtz free energy per unit surface given by Eq. 56 into Eq. 36, the interfacial recoverable microstresses $\mathbb{M}_{ij}^{I,\text{en}}$ can be obtained as follows:

$$\mathbb{M}_{ij}^{I,\text{en}} = G \ell_{\text{en}}^I \varepsilon_{ij}^{pI(\text{pre})} \quad (57)$$

As can be seen in Eq. 57, $\mathbb{M}_{ij}^{I,\text{en}}$ does not involve the plastic strain rate, which is related to the dislocation slip, and the temperature since the interfacial recoverable microstresses are activated by the recoverable stored energy.

Dissipation Potential and Dissipative Thermodynamic Microstresses

In this section, the dissipation potential functions for the aforementioned two cases of the SGP model are postulated, respectively. The first case is derived from the dissipation potential dependent on the plastic strain gradient, which leads to the nonzero dissipative thermodynamic microstress $\mathcal{S}_{ijk}^{\text{dis}} \neq 0$, while the other case is derived from the dissipation potential that is independent on the plastic strain gradient, which leads to $\mathcal{S}_{ijk}^{\text{dis}} = 0$.

Coleman and Gurtin (1967) pointed out that the dissipation potential function is composed of two parts, the mechanical part which is dependent on the plastic strain and its gradient and the thermal counterpart which shows the purely thermal effect such as the heat conduction. In this sense, and in the context of Eq. 37, the functional form of the dissipation potential, which is dependent on $\dot{\varepsilon}_{ij,k}^p$, for the former class can be put forward as (Voyiadjis and Song 2017):

$$\mathcal{D} = \frac{\sigma_y}{2\dot{\rho}} \left\{ \left(\frac{\dot{\rho}}{\dot{\rho}_0} \right)^m \left(1 - \left(\frac{T}{T_y} \right)^n \right) \left(\dot{\varepsilon}_{ij}^p \dot{\varepsilon}_{ij}^p \right) + \ell_{\text{dis}}^2 \left(\frac{\dot{\rho}}{\dot{\rho}_0} \right)^m \left(1 - \left(\frac{T}{T_y} \right)^n \right) \left(\dot{\varepsilon}_{\text{lm},k}^p \dot{\varepsilon}_{\text{lm},k}^p \right) \right\} - \frac{b}{2} \dot{T}^2 - \frac{1}{2} \frac{k(T)}{T} T_i T_i \geq 0 \quad (58)$$

where σ_y is a material constant accounting for the yield strength, m is a nonnegative material constant for the rate sensitivity parameter, in which the limit $m \rightarrow 0$ corresponds to rate-independent material behavior, $\dot{\rho}_0$ is a constant for the reference flow rate, ℓ_{dis} is the dissipative length scale that corresponds to the dissipative effects in terms of the gradient of the plastic strain rate, b is the material constant accounting for the energy exchange between phonon and electron, and $k(T)$ is the thermal conductivity coefficient. The generalized dissipative effective plastic strain measure $\dot{\rho}$ is defined as a function of the plastic strain rate, the gradient of the plastic strain

rate, and the dissipative length scale as follows:

$$\dot{\rho} = \sqrt{\dot{\epsilon}_{ij}^p \dot{\epsilon}_{ij}^p + \ell_{\text{dis}}^2 \dot{\epsilon}_{\text{lm},k}^p \dot{\epsilon}_{\text{lm},k}^p} \tag{59}$$

By using the dissipation potential given in Eq. 58 along with Eq. 39 and the assumption $k(T)/T = k_0 = \text{constant}$, the dissipative thermodynamic forces for the former case (Case I) can be obtained as follows:

$$\mathcal{X}_{ij}^{\text{dis}} = \sigma_y \left(\frac{\dot{\rho}}{\dot{\rho}_0} \right)^m \left(1 - \left(\frac{T}{T_y} \right)^n \right) \frac{\dot{\epsilon}_{ij}^p}{\dot{\rho}} \tag{60}$$

$$\mathcal{S}_{ijk}^{\text{dis}} = \sigma_y \ell_{\text{dis}}^2 \left(\frac{\dot{\rho}}{\dot{\rho}_0} \right)^m \left(1 - \left(\frac{T}{T_y} \right)^n \right) \frac{\dot{\epsilon}_{ij,k}^p}{\dot{\rho}} \tag{61}$$

$$\mathcal{A}^{\text{dis}} = -b\dot{T} \tag{62}$$

$$q_i = -k_0 T_{,i} \tag{63}$$

On the other hand, the functional form of the dissipation potential, which is independent of $\dot{\epsilon}_{ij,k}^p$, for the latter case (Case II) can be postulated by setting $\ell_{\text{dis}} = 0$ in Eq. 58 as follows:

$$\mathcal{D} = \frac{\sigma_y}{2} \left(\frac{\dot{\rho}_{\ell_{\text{dis}}=0}}{\dot{\rho}_0} \right)^m \left(1 - \left(\frac{T}{T_y} \right)^n \right) \dot{\rho}_{\ell_{\text{dis}}=0} - \frac{b}{2} \dot{T}^2 - \frac{1}{2} \frac{k(T)}{T} T_{,i} T_{,i} \geq 0 \tag{64}$$

where $\dot{\rho}_{\ell_{\text{dis}}=0}$ is given by $\sqrt{\dot{\epsilon}_{ij}^p \dot{\epsilon}_{ij}^p}$ by setting $\ell_{\text{dis}} = 0$ in Eq. 59. By substituting Eq. 64 into Eq. 39, the dissipative thermodynamic forces for the latter case (Case II) can be obtained as follows:

$$\mathcal{X}_{ij}^{\text{dis}} = \sigma_y \left(\frac{\dot{\rho}_{\ell_{\text{dis}}=0}}{\dot{\rho}_0} \right)^m \left(1 - \left(\frac{T}{T_y} \right)^n \right) \frac{\dot{\epsilon}_{ij}^p}{\dot{\rho}_{\ell_{\text{dis}}=0}} \tag{65}$$

$$\mathcal{S}_{ijk}^{\text{dis}} = 0 \tag{66}$$

$$\mathcal{A}^{\text{dis}} = -b\dot{T} \tag{67}$$

$$q_i = -k_0 T_{,i} \tag{68}$$

Meanwhile, Gurtin and Reddy (2009) pointed out that the classical isotropic hardening rule, which is dissipative in nature, may equally well be characterized via a defect energy since this energetically based hardening rule mimics the dissipative behavior by describing loading processes that are irreversible. In this

sense, as it was mentioned previously in this chapter, the energetic microstress $\mathcal{X}_{ij}^{\text{en}}$ is further decomposed into $\mathcal{X}_{ij}^{\text{en,NR}}$ that describes an irreversible loading process and $\mathcal{X}_{ij}^{\text{en,R}}$ that describes a reversible loading process due to the energy carried by the dislocations. The present framework follows the theorem of Gurtin and Reddy (2009) in that the theory without a defect energy is equivalent to the theory with a defect energy by replacing the dissipation \mathcal{D} by an effective dissipation \mathcal{D}_{eff} , which is defined as follows:

$$\mathcal{D}_{\text{eff}} = \left(\mathcal{X}_{ij}^{\text{dis}} + \mathcal{X}_{ij}^{\text{en,NR}} \right) \dot{\varepsilon}_{ij}^p + \mathcal{S}_{ijk}^{\text{dis}} \dot{\varepsilon}_{ij,k}^p + \mathcal{A}^{\text{dis}} \dot{T} - \frac{q_i}{T} T_{,i} \quad (69)$$

where $\mathcal{X}_{ij}^{\text{en,NR}}$ may be viewed as the effectively dissipative microforce since it satisfies an effective dissipation inequality.

There are two main mechanisms affecting the energy dissipation during the dislocation movement in the grain boundary area. The first mechanism is related to an energy change in the grain boundary region. The macroscopic accumulated plastic strain at the grain boundary can be connected to the microscopic deformation of the grain boundary through the quadratic mean of the deformation gradient. Thus, the energy change after the onset of slip transmission to the adjacent grain is able to be approximately determined by a quadratic function of the deformation gradient at the microscale and hence the interfacial plastic strain at the macroscale. The other mechanism introduces the energy involved in the deformation of the grain boundary. This energy is mainly due to the energy dissipation during the dislocation movement and can be taken as a linear function of the interfacial plastic strain.

The interfacial dissipation potential \mathcal{D}^I in the current study is postulated by combing the above-mentioned mechanisms as follows:

$$\mathcal{D}^I = \frac{\ell_{\text{dis}}^I}{m^I + 1} \left(\sigma_y^I + h^I \varepsilon_p^{I(\text{post})} \right) \left(1 - \frac{T^I}{T_y^I} \right)^{n^I} \left(\frac{\dot{\varepsilon}_p^{I(\text{post})}}{\dot{\varepsilon}_0^I} \right)^{m^I} \dot{\varepsilon}_p^{I(\text{post})} \geq 0 \quad (70)$$

where ℓ_{dis}^I is the interfacial dissipative length scale, m^I and $\dot{\varepsilon}_0^I$ are the viscous related material parameters, σ_y^I is a constant accounting for the interfacial yield stress at which the interface starts to deform plastically, h^I is an interfacial hardening parameter representing the slip transmission through the interface, T_y^I is the scale-independent interfacial thermal parameter at the onset of yield, n^I is the interfacial thermal parameter, and $\varepsilon_p^{I(\text{post})} = \sqrt{\varepsilon_{ij}^{pI(\text{post})} \varepsilon_{ij}^{pI(\text{post})}}$ and $\dot{\varepsilon}_p^{I(\text{post})} = \sqrt{\dot{\varepsilon}_{ij}^{pI(\text{post})} \dot{\varepsilon}_{ij}^{pI(\text{post})}}$ are defined, respectively, with the plastic strain at the interface after the slip transfer $\varepsilon_{ij}^{pI(\text{post})}$ and its rate $\dot{\varepsilon}_{ij}^{pI(\text{post})}$. The rate-dependency and temperature-dependency of the interfacial dissipation energy are clearly shown in Eq. 70 through the terms $\left(\dot{\varepsilon}_p^{I(\text{post})} / \dot{\varepsilon}_0^I \right)^{m^I}$ and $\left(1 - T^I / T_y^I \right)^{n^I}$, respectively.

The interfacial dissipative microstresses $\mathbb{M}_{ij}^{I,\text{dis}}$ can be obtained by substituting Eq. 70 into Eq. 40 as follows:

$$\mathbb{M}_{ij}^{I,\text{dis}} = \frac{\ell_{\text{dis}}^I}{m^I + 1} \left(\sigma_y^I + h^I \varepsilon_p^{I(\text{post})} \right) \left(1 - \frac{T^I}{T_y^I} \right)^{n^I} \left(\frac{\dot{\varepsilon}_p^{I(\text{post})}}{\dot{\gamma}_0^I} \right)^{m^I} \frac{\dot{\varepsilon}_{ij}^{pI(\text{post})}}{\dot{\varepsilon}_p^{I(\text{post})}} \quad (71)$$

By substituting Eqs. 71 and 57 into Eq. 31, one can obtain the interfacial microtraction \mathbb{M}_{ij}^I as follows:

$$\begin{aligned} \mathbb{M}_{ij}^I &= G \ell_{\text{en}}^I \varepsilon_{ij}^{pI(\text{pre})} \\ &+ \frac{\ell_{\text{dis}}^I}{m^I + 1} \left(\sigma_y^I + h^I \varepsilon_p^{I(\text{post})} \right) \left(1 - \frac{T^I}{T_y^I} \right)^{n^I} \left(\frac{\dot{\varepsilon}_p^{I(\text{post})}}{\dot{\gamma}_0^I} \right)^{m^I} \frac{\dot{\varepsilon}_{ij}^{pI(\text{post})}}{\dot{\varepsilon}_p^{I(\text{post})}} \end{aligned} \quad (72)$$

As can be seen in Eq. 72, a free surface, i.e., microfree boundary condition, at the grain boundary can be described by setting $\ell_{\text{en}}^I = \ell_{\text{dis}}^I = 0$ and it is also possible to describe a surface passivation, i.e., microclamped boundary condition, by setting $\ell_{\text{en}}^I \rightarrow \infty$ and $\ell_{\text{dis}}^I \rightarrow \infty$.

Flow Rule

The flow rule in the present framework is established based on the nonlocal microforce balance, Eq. 6, augmented by thermodynamically consistent constitutive relations for both energetic and dissipative microstresses. By substituting Eqs. 51, 53, 60, and 61 into Eq. 6, one can obtain a second-order partial differential form of the flow rule as follows (The flow rule, Eq. 73, corresponds to the SGP model (Case I), in which the functional form of the dissipation potential is dependent on $\dot{\varepsilon}_{ij,k}^p$ ($S_{ijk}^{\text{dis}} \neq 0$)). One can easily obtain the flow rule for the other case (Case II) of the SGP model, i.e. with $S_{ijk}^{\text{dis}} = 0$, by setting $\ell_{\text{dis}} = 0$.) (Voyiadjis and Song 2017):

$$\begin{aligned} \tau_{ij} &- \underbrace{\left(-G \ell_{\text{en}}^2 \varepsilon_{ij,kk}^p \right)}_{-S_{ijk,k}^{\text{en}}} \\ &= h \varepsilon_{ij}^p \left(1 - \left(\frac{T}{T_y} \right)^n \right) \varepsilon_p^{r-1} + \sigma_y \left(\frac{\dot{\gamma}}{\dot{\gamma}_0} \right)^m \left(1 - \left(\frac{T}{T_y} \right)^n \right) \frac{\dot{\varepsilon}_{ij}^p}{\dot{\gamma}} \\ &- \left(\sigma_y \ell_{\text{dis}}^2 \left(\frac{\dot{\gamma}}{\dot{\gamma}_0} \right)^m \left(1 - \left(\frac{T}{T_y} \right)^n \right) \frac{\dot{\varepsilon}_{ij,kk}^p}{\dot{\gamma}} \right) \end{aligned} \quad (73)$$

where the under-braced term $S_{ijk,k}^{\text{en}}$ represents a backstress due to the energy stored in dislocations and results in the Bauschinger effect observed in the experiments

(Liu et al. 2015; Nicola et al. 2006; Xiang et al. 2005; Xiang and Vlassak 2006) and discrete dislocation plasticity (Nicola et al. 2006; Shishvan et al. 2010, 2011).

Thermo-mechanical Coupled Heat Equation

The evolution of the temperature field is governed by the law of conservation of energy (the first law of thermodynamics). The terms addressing heating as a result of the inelastic dissipation and thermo-mechanical coupling are involved for describing the evolution of the temperature field. The equation for the conservation of energy in this chapter is put forward as follows:

$$\sigma_{ij} \dot{\varepsilon}_{ij}^e + \mathcal{X}_{ij} \dot{\varepsilon}_{ij}^p + \mathcal{S}_{ijk} \dot{\varepsilon}_{ij,k}^p + \mathcal{A} \dot{T} + \mathcal{B}_i \dot{T}_{,i} - \text{div } q_i + \rho \mathcal{H}_{\text{EXT}} - \rho \dot{e} = 0 \quad (74)$$

where \mathcal{H}_{EXT} is the specific heat from the external source.

By considering the effective dissipation potential given in Eq. 69 along with the equations for the entropy production (the second law of thermodynamics) previously described in Eq. 17, the relationship for the evolution of the entropy, which describes the irreversible process, can be derived as follows:

$$\rho \dot{s} T = \mathcal{D}_{eff} + \rho \mathcal{H}_{\text{EXT}} \quad (75)$$

By using Eq. 50 for solving the rate of the entropy \dot{s} , the evolution of the temperature field can be obtained as follows:

$$\begin{aligned} \rho c_0 \dot{T} = & \underbrace{\mathcal{X}_{ij}^{\text{dis}} \dot{\varepsilon}_{ij}^p + \mathcal{X}_{ij}^{\text{en,IR}} \dot{\varepsilon}_{ij}^p + \mathcal{S}_{ijk}^{\text{dis}} \dot{\varepsilon}_{ij,k}^p}_{\textcircled{1}} + \underbrace{\frac{k(T)}{2T} T_{,i} T_{,i} + \alpha T_r \dot{T}_{,ii} - \mathbf{b} T_r \ddot{T}}_{\textcircled{2}} \\ & - \underbrace{\alpha_t \dot{\varepsilon}_{ij}^e \delta_{ij} T}_{\textcircled{3}} \\ & - \underbrace{\left\{ h \varepsilon_p^{r-1} \left(\frac{n}{T_y} \left(\frac{T}{T_y} \right)^{n-1} \right) \dot{\varepsilon}_{ij}^p + \frac{h \varepsilon_p^{r+1}}{r+1} \left(\frac{n(n-1)}{T_y^2} \left(\frac{T}{T_y} \right)^{n-2} \right) \dot{T} \right\} T}_{\textcircled{4}} \\ & + \underbrace{\rho \mathcal{H}_{\text{EXT}}}_{\textcircled{5}} \end{aligned} \quad (76)$$

where c_0 is the specific heat capacity at constant volume and is given by $c_0 = \text{constant} \cong c_e T/T_r$. As shown in Eq. 76, the following terms are depicted, $\textcircled{1}$ irreversible mechanical process, $\textcircled{2}$ generalized heat conduction, $\textcircled{3}$ thermo-elastic coupling, $\textcircled{4}$ thermo-plastic coupling, and $\textcircled{5}$ heat source, which are involved in the

evolution of the temperature. The third-order mixed derivation term $\alpha T_r T_{,ii}$ in part ② introduces the microstructural interaction effect into the classical heat equation, in addition, the second-order time derivative term $b T_r \ddot{T}$ gives the thermal wave behavior effect in heat propagation.

By substituting the constitutive equations of the energetic microstresses given by Eqs. 49, 50, 51, 52, and 53 and the dissipative microstresses given by Eqs. 60, 61, 62, and 63 into Eq. 76 and defining three additional terms $\ell_{\text{eff}} = k_0/\rho c_v$, $\ell^{\text{en}} = \alpha T_r/k_0$, and $\ell^{\text{dis}} = b T_r/\rho c_v$, the evolution of temperature for the model, in which the functional form of the dissipation potential dependent on $\dot{\varepsilon}_{ij,k}^p$, i.e., $S_{ijk}^{\text{dis}} \neq 0$, can be obtained as follows (Voyiadjis and Song 2017):

$$\begin{aligned} & \left[1 + \frac{h \varepsilon_p^{r+1}}{\rho c_v (r+1)} \left(\frac{n(n-1)}{T_y} \left(\frac{T}{T_y} \right)^n \right) \right] \dot{T} \\ &= \frac{\dot{\varepsilon}_{ij}^p}{\rho c_v} \left[\left\{ h \varepsilon_p^{r-1} \left(1 - \left(\frac{T}{T_y} \right)^n \right) - 2r \vartheta h^2 \varepsilon_p^{2r-1} - n h \varepsilon_p^{r-1} \left(\frac{T}{T_y} \right)^n \right\} \varepsilon_{ij}^p \right. \\ & \quad \left. + \sigma_y \left(\frac{\dot{\rho}}{\dot{\rho}_0} \right)^m \left(1 - \left(\frac{T}{T_y} \right)^n \right) \frac{\dot{\varepsilon}_{ij}^p}{\dot{\rho}} \right] \\ & \quad + \frac{1}{\rho c_v} \left[\sigma_y \ell_{\text{dis}}^2 \left(\frac{\dot{\rho}}{\dot{\rho}_0} \right)^m \left(1 - \left(\frac{T}{T_y} \right)^n \right) \frac{\dot{\varepsilon}_{ij,k}^p \dot{\varepsilon}_{ij,k}^p}{\dot{\rho}} - \alpha_i \dot{\varepsilon}_{ij}^e \delta_{ij} T + \rho \mathcal{H}_{\text{EXT}} \right] \\ & \quad + [\ell_{\text{eff}} T_{,ii} + \ell_{\text{eff}} \ell^{\text{en}} \dot{T}_{,ii} - \ell^{\text{dis}} \ddot{T}] \end{aligned} \quad (77)$$

In the absence of the mechanical terms, Eq. 77 turns to the generalized heat equation including ℓ^{en} and ℓ^{dis} . The evolution of temperature for the other case (Case II) of SGP model, i.e., with $S_{ijk}^{\text{dis}} = 0$ can be obtained by setting $\ell_{\text{dis}} = 0$ in Eq. 77.

Finite Element Implementation of the Strain Gradient Plasticity Model

In this section, first a one-dimensional finite element model for the SGP model by Voyiadjis and Song (2017) is presented to investigate the size dependent behavior in the microscopic structures under macroscopically uniform uniaxial tensile stress. In a one-dimensional finite element implementation, the macroscopic partial differential equations for balance of linear momentum Eq. 5 with the macroscopic boundary conditions, $u_{x=0} = 0$ and $u_{x=L} = u^\dagger$ (prescribed), and the microscopic partial differential equations for nonlocal force balance Eq. 6 with the microscopic boundary conditions, $(\mathbb{M}^I - \mathcal{S})_{x=0} = 0$ and $(\mathbb{M}^I - \mathcal{S})_{x=L} = 0$, yield the following expressions in a global weak form, respectively (The finite element solutions, in this section, depend on the x -direction. A single crystal with the size

of L bounded by two grain boundaries is analyzed (see Fig. 7 for details). Also a two dimensional case is presented for solving the problem of the square plate in the latter part of this chapter:

$$\int_0^L (\sigma \tilde{u}_{,x}) dx = 0 \quad (78)$$

$$\int_0^L [(\mathcal{X} - \tau) \tilde{\varepsilon}^p + \mathcal{S} \tilde{\varepsilon}_{,x}^p] dx + \mathbb{M}^I \tilde{\varepsilon}_{x=L}^p - \mathbb{M}^I \tilde{\varepsilon}_{x=0}^p = 0 \quad (79)$$

where the arbitrary virtual fields \tilde{u} and $\tilde{\varepsilon}^p$ are assumed to be kinematically admissible weighting functions in the sense that $\tilde{u}_{x=0} = \tilde{u}_{x=L} = 0$. (In the case of micro-clamped boundary condition, $\tilde{\varepsilon}_{x=0}^p = \tilde{\varepsilon}_{x=L}^p = 0$ is imposed at the grain boundaries to enforce the complete blockage of dislocations at the interface. In the case of micro-free boundary condition, on the other hand, the dislocations are free to move across the interface, which in turn, the present grain boundary flow rule is imposed.)

The user-element subroutine UEL in the commercial finite element package ABAQUS/standard (2012) is presented in this chapter in order to numerically solve the weak forms of the macroscopic and microscopic force balances, Eqs. 78 and 79, respectively. In this finite element solution, the displacement field u and the plastic strain field ε^p are discretized independently and both of the fields are taken as fundamental unknown nodal degrees of freedom. In this regard, the displacement field and corresponding strain field ε , and the plastic strain field and corresponding plastic strain gradient field $\varepsilon_{,x}^p$ can be obtained by using the interpolation as follows:

$$u(x) = \sum_{\xi=1}^{n_u} \mathbb{N}_u^\xi \mathcal{U}_u^\xi \varepsilon(x) = \frac{\partial u(x)}{\partial x} = \sum_{\xi=1}^{n_u} \mathbb{N}_{u,x}^\xi \mathcal{U}_u^\xi \quad (80)$$

$$\varepsilon^p(x) = \sum_{\xi=1}^{n_{\varepsilon^p}} \mathbb{N}_{\varepsilon^p}^\xi \mathcal{E}_{\varepsilon^p}^\xi \varepsilon_{,x}^p(x) = \frac{\partial \varepsilon^p(x)}{\partial x} = \sum_{\xi=1}^{n_{\varepsilon^p}} \mathbb{N}_{\varepsilon^p,x}^\xi \mathcal{E}_{\varepsilon^p}^\xi \quad (81)$$

where \mathbb{N}_u^ξ and $\mathbb{N}_{\varepsilon^p}^\xi$ are the shape functions, and \mathcal{U}_u^ξ and $\mathcal{E}_{\varepsilon^p}^\xi$ are the nodal values of the displacements and the plastic strains at node ξ , respectively. The terms n_u and n_{ε^p} represent the number of nodes per a single element for the displacement and the plastic strain, respectively. (If a one-dimensional three-noded quadratic element is employed, n_u and n_{ε^p} are set up as three. On the other hand, these parameters are set up as two in the case that a one-dimensional two-noded linear element is used. It should be noted that n_u and n_{ε^p} do not necessarily have to be same as each other in the present finite element implementation, even though both the displacement field and the plastic strain fields are calculated by using the standard isoparametric shape functions.)

Substituting Eqs. 80 and 81 into Eqs. 78 and 79 give the nodal residuals for the displacement r_u and the plastic strain r_{ε^p} for each finite element el as follows:

$$(r_u)_\xi = \int_{el} (\sigma N_{u,x}^\xi) dx \tag{82}$$

$$(r_{\varepsilon^p})_\xi = \int_{el} [(\mathcal{X} - \tau) N_{\varepsilon^p}^\xi + \mathcal{S} N_{\varepsilon^p,x}^\xi] dx + M^I N_{\varepsilon^p}^\xi \tag{83}$$

where the term $M^I N_{\varepsilon^p}^\xi$ is applied only for the nodes on the interface which is at $x = 0$ and $x = L$.

The global coupled system of equations, $(r_u)_\xi = 0$ and $(r_{\varepsilon^p})_\xi = 0$, are solved using ABAQUS/standard (2012) based on the Newton-Raphson iterative scheme. Occasionally, the modified Newton-Raphson method, referred to as quasi Newton-Raphson method, is employed in the case that the numerical solution suffers a divergence during the initial increment immediately after an abrupt change in loading. The Taylor expansion of the residuals with regard to the current nodal values can be expressed by assuming the nodal displacement and the plastic strain in iteration ζ as U_u^ζ and $\mathcal{E}_{\varepsilon^p}^\zeta$, respectively, as follows:

$$\begin{aligned} (r_u|_{U_u^{\zeta+1}, \mathcal{E}_{\varepsilon^p}^{\zeta+1}})_\xi &= (r_u|_{U_u^\zeta, \mathcal{E}_{\varepsilon^p}^\zeta})_\xi + \left(\frac{\partial r_u}{\partial U_u^\zeta} \Big|_{U_u^\zeta} \right) \Delta U_u^\xi + \left(\frac{\partial r_u}{\partial \mathcal{E}_{\varepsilon^p}^\zeta} \Big|_{\mathcal{E}_{\varepsilon^p}^\zeta} \right) \Delta \mathcal{E}_{\varepsilon^p}^\xi \\ &+ \mathcal{O} \left((\Delta U_u^\xi)^2, (\Delta \mathcal{E}_{\varepsilon^p}^\xi)^2 \right) \end{aligned} \tag{84}$$

$$\begin{aligned} (r_{\varepsilon^p}|_{U_u^{\zeta+1}, \mathcal{E}_{\varepsilon^p}^{\zeta+1}})_\xi &= (r_{\varepsilon^p}|_{U_u^\zeta, \mathcal{E}_{\varepsilon^p}^\zeta})_\xi + \left(\frac{\partial r_{\varepsilon^p}}{\partial U_u^\zeta} \Big|_{U_u^\zeta} \right) \Delta U_u^\xi + \left(\frac{\partial r_{\varepsilon^p}}{\partial \mathcal{E}_{\varepsilon^p}^\zeta} \Big|_{\mathcal{E}_{\varepsilon^p}^\zeta} \right) \Delta \mathcal{E}_{\varepsilon^p}^\xi \\ &+ \mathcal{O} \left((\Delta U_u^\xi)^2, (\Delta \mathcal{E}_{\varepsilon^p}^\xi)^2 \right) \end{aligned} \tag{85}$$

where $\Delta U_u^\xi = (U_u^{\zeta+1})_\xi - (U_u^\zeta)_\xi$, $\Delta \mathcal{E}_{\varepsilon^p}^\xi = (\mathcal{E}_{\varepsilon^p}^{\zeta+1})_\xi - (\mathcal{E}_{\varepsilon^p}^\zeta)_\xi$, and $\mathcal{O} \left((\Delta U_u^\xi)^2, (\Delta \mathcal{E}_{\varepsilon^p}^\xi)^2 \right)$ are the big \mathcal{O} notation to represent the terms of higher order than the second degree. The residual is ordinarily calculated at the end of each time step, and the values of the nodal displacements and the plastic strains are updated during the iterations. The increments in nodal displacements and the plastic strains can be computed by solving the system of linear equations shown in Eq. 86 with the Newton-Raphson iterative method:

$$\underbrace{\begin{bmatrix} K_{uu}^{el} & K_{u\varepsilon^p}^{el} \\ K_{\varepsilon^p u}^{el} & K_{\varepsilon^p \varepsilon^p}^{el} \end{bmatrix}}_{K^{el}} \begin{Bmatrix} \Delta U_u^\xi \\ \Delta \mathcal{E}_{\varepsilon^p}^\xi \end{Bmatrix} = \begin{Bmatrix} (r_u|_{U_u^\zeta, \mathcal{E}_{\varepsilon^p}^\zeta})_\xi \\ (r_{\varepsilon^p}|_{U_u^\zeta, \mathcal{E}_{\varepsilon^p}^\zeta})_\xi \end{Bmatrix} \tag{86}$$

where K^{el} is the Jacobian (stiffness) matrix that needs to be defined in the user-subroutine for each element.

From Eqs. 84 and 85 along with the discretization for the displacements given by Eq. 80 and the plastic strains given by Eq. 81 at the end of a time step, and the functional forms of the energetic and dissipative higher order stresses defined in the previous sections, the Jacobian matrix for the case of the SGP model with $S_{ijk}^{dis} \neq 0$ can be obtained as follows:

$$K_{uu}^{el} = - \frac{\partial r_u}{\partial \mathcal{U}_u^\xi} \Big|_{\mathcal{U}_u^\xi} = - \int_{el} (E N_{u,x}^\xi N_{u,x}^\xi) dx \quad (87)$$

$$K_{ue^p}^{el} = - \frac{\partial r_{e^p}}{\partial \mathcal{E}_{e^p}^\xi} \Big|_{\mathcal{E}_{e^p}^\xi} = \int_{el} (E N_{e^p}^\xi N_{u,x}^\xi) dx \quad (88)$$

$$K_{e^p u}^{el} = - \frac{\partial r_{e^p}}{\partial \mathcal{U}_u^\xi} \Big|_{\mathcal{U}_u^\xi} = \int_{el} (E N_{u,x}^\xi N_{e^p}^\xi) dx \quad (89)$$

$$\begin{aligned} (K_{e^p e^p}^{el})_{S_{ijk}^{dis} \neq 0} &= \left(- \frac{\partial r_{e^p}}{\partial \mathcal{E}_{e^p}^\xi} \Big|_{\mathcal{E}_{e^p}^\xi} \right)_{S_{ijk}^{dis} \neq 0} \\ &= \int_{el} \left[\underbrace{\left\{ E + \left(\text{rh}(\varepsilon^p)^r + \frac{\sigma_y (\dot{\gamma})^{m-1}}{\Delta t (\dot{\gamma}_0)^m} \right) \left(1 - \left(\frac{T}{T_y} \right)^n \right) \right\}}_{\text{for bulk}} N_{e^p}^\xi N_{e^p}^\xi \right. \\ &\quad + \left. \left\{ \left(G \ell_{en}^2 + \frac{\sigma_y \ell_{dis}^2 (\dot{\gamma})^{m-1}}{\Delta t (\dot{\gamma}_0)^m} \right) \left(1 - \left(\frac{T}{T_y} \right)^n \right) \right\}}_{\text{for bulk}} N_{e^p,x}^\xi N_{e^p,x}^\xi \right] dx \\ &\quad + \left[\underbrace{\left(G (\ell_{en}^I)^2 + \frac{\sigma_y^I (\ell_{dis}^I)^2 (\dot{\gamma}^I)^{m^I-1}}{\Delta t (\dot{\gamma}_0^I)^{m^I}} \right) \left(1 - \left(\frac{T}{T_y^I} \right)^{n^I} \right)}_{\text{for interface}} \right] N_{e^p}^\xi N_{e^p}^\xi \end{aligned} \quad (90)$$

where Δt is a time step. The Jacobian matrix for the other case, i.e. $(K_{e^p e^p}^{el})_{S_{ijk}^{dis} = 0}$, can be obtained by setting $\ell_{dis} = 0$ in Eq. 90. The interfacial terms in $(K_{e^p e^p}^{el})_{S_{ijk}^{dis} \neq 0}$ and $(K_{e^p e^p}^{el})_{S_{ijk}^{dis} = 0}$ are applied only for the nodes on the interface which is at $x = 0$ and $x = L$ in this chapter.

Experimental Validation of the Strain Gradient Plasticity Model

In this section, the present SGP model and corresponding finite element code by Voyiadjis and Song (2017) are validated by comparing with the experimental results from three sets of size effect experiments. In addition, the comparison between the

SGP model by Voyiadjis and Song (2017) and Voyiadjis and Faghihi (2014) is carried out to show the increase in accuracy of the model by Voyiadjis and Song (2017). To examine the applicability of the present finite element implementation to the various kinds of materials, it is considered that each set of three experiments involves the three different materials, viz. aluminum (Al), copper (Cu), and nickel (Ni).

Haque and Saif (2003) developed Micro-Electro Mechanical Systems (MEMS)-based testing techniques for uniaxial tensile testing of nanoscale freestanding Al thin films to explore the effect of strain gradient in 100 nm, 150 nm, and 485 nm thick specimens with average grain size of 50 nm, 65 nm, and 212 nm, respectively. The specimens with 99.99% pure sputter-deposited freestanding Al thin films are 10 μm wide and 275 μm long. All experiments are carried out in situ in SEM and the stress and strain resolutions for the tests are set 5 Mpa and 0.03%, respectively. In particular, the comparison between the present SGP model by Voyiadjis and Song (2017) and Voyiadjis and Faghihi (2014) is carried out to show the increase in accuracy of the present model. The calibrated material parameters as well as the general material parameters for the Al are presented in Table 1, and the numerical results from both the present model by Voyiadjis and Song (2017) and Voyiadjis and Faghihi (2014) are shown in Fig. 2 in conjunction with the experimental data of Haque and Saif (2003). As it is clearly shown in this figure, the size effect: Smaller is Stronger is observed on the stress-strain curves of the Al thin films. Furthermore, the calculated results of the present SGP model by Voyiadjis and Song (2017) display a

Table 1 The general and calibrated material parameters used for the validation of the proposed strain gradient plasticity model (Reprinted with permission from Voyiadjis and Song 2017)

General		Aluminum	Copper	Nickel
E (GPa)	Elastic modulus for isotropic linear elasticity	110	70	115
G (GPa)	Shear modulus for isotropic linear elasticity	48	27	44
ρ (g cm^{-3})	Density	8.960	2.702	8.902
c_3 (J/g K)	Specific heat capacity at constant stress	0.385	0.910	0.540
Calibrated		Aluminum	Copper	Nickel
σ_y (MPa)	Yield stress	195	700	950
h (MPa)	Hardening material parameter	600	1,700	3,500
$\dot{\epsilon}_0$ (s^{-1})	Reference effective plastic strain rate	0.04	0.04	0.04
r	Nonlinear hardening material parameter	0.6	0.2	0.2
m	Non-negative rate sensitivity parameter	0.05	0.05	0.05
T_y (K)	Thermal material parameter	1,358	933	890
n	Temperature sensitivity parameter	0.3	0.3	0.3
ℓ_{en} (μm)	Energetic length scale	1.5 (1.0 μm)	0.9 (100 nm)	1.0
ℓ_{dis} (μm)	Dissipative length scale	2.5(1.0 μm)	8.0 (100 nm)	0.1

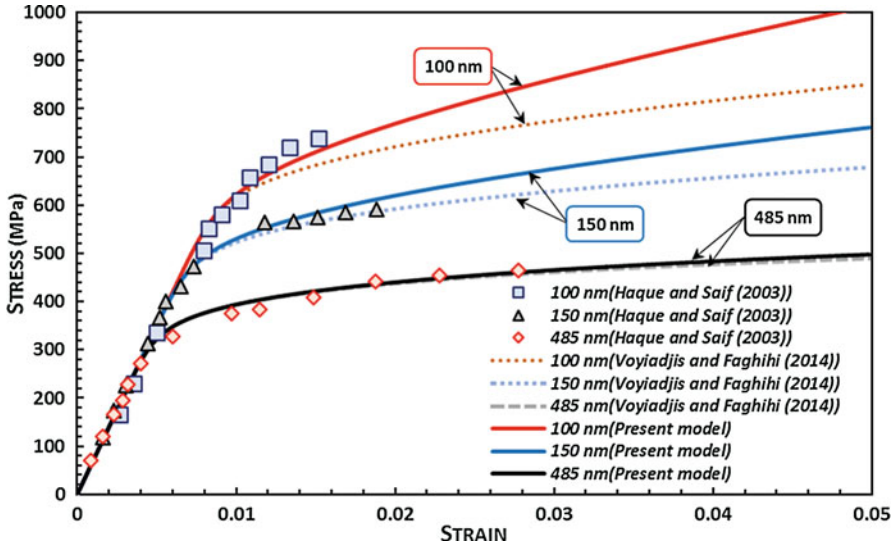


Fig. 2 The validation of the present strain gradient plasticity model by comparing the numerical results from the present model by Voyiadjis and Song (2017) with those from Voyiadjis and Faghihi (2014) and the experimental measurements from Haque and Saif (2003) on the stress-strain response of the sputter-deposited Al thin films (Reprinted with permission from Voyiadjis and Song 2017)

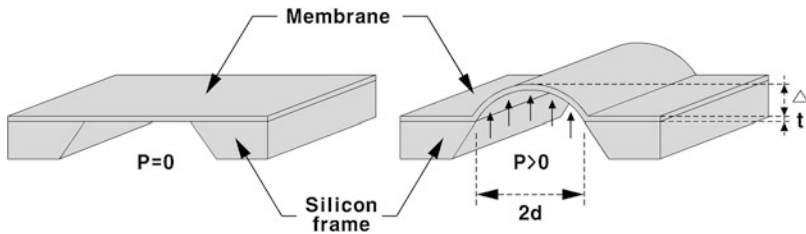


Fig. 3 Schematic representation of the plane strain bulge test technique presented by Xiang and Vlassak (2006). The stress-strain response of either a single material membrane or a stack of multiple material membranes adhered on a rigid Si frame is able to be obtained by using the following equations: $\sigma = P(d^2 + \Delta^2)/2t\Delta$ and $\epsilon = \epsilon_r + \{(d^2 + \Delta^2)/2d\Delta\} \arcsin(2d\Delta/d^2 + \Delta^2) - 1$ where ϵ_r is a residual strain in the membrane (Reprinted with permission from Voyiadjis and Song 2017)

tendency to be more coincident to the experimental data than those of Voyiadjis and Faghihi (2014).

Xiang and Vlassak (2006) investigated the size effects with a variety of film thicknesses on the plastic behavior of the freestanding electroplated Cu thin films by performing the plane strain bulge test. In this plane strain bulge test, the rectangular freestanding membranes surrounded by a rigid silicon (Si) frame are deformed in plane strain by applying a uniform pressure to one side of the membrane as shown in Fig. 3. The displacement and pressure resolutions for this bulge tests system are 0.3 μm and 0.1 kpa, respectively.

As can be seen in the work of Xiang and Vlassak (2006), the stress-strain curves of the Cu thin films with a passivation layer on both surfaces clearly show the size effects due to the presence of a boundary layer with high dislocation density near the film-passivation layer interfaces. In this sense, the bulge test of electroplated Cu thin films with both surfaces passivated by 20 nm of titanium (Ti) is considered here for the experimental validation of the present SGP model. In order to describe the passivation effect, the microclamped condition, which causes the dislocations to be completely blocked at the grain boundary, is imposed at both surfaces of the Cu thin films. Meanwhile, the experiments are performed with the various thicknesses of the Cu thin films of 1.0 μm , 1.9 μm , and 4.2 μm . The average grain sizes in all cases are given by $1.5 \pm 0.05 \mu\text{m}$, $1.51 \pm 0.04 \mu\text{m}$, and $1.5 \pm 0.05 \mu\text{m}$, respectively, which mean almost equal to each other.

The calibrated and general material parameters for the copper are presented in Table 1, and the comparison between the experimental measurements from the bulge tests and the calculated results from the present SGP model by Voyiadjis and Song (2017) is shown in Fig. 4. As it is clearly shown in this figure, the size effects according to the variation of the Cu thin film thicknesses is well observed in both the present SGP model and the experimental work of Xiang and Vlassak (2006). Moreover, the numerical results of the present model by Voyiadjis and Song (2017) are in good agreement with the experimental measurements.

Han et al. (2008) developed the microscale tensile testing system, which is composed of a high temperature furnace, a micro motor actuator and the Digital

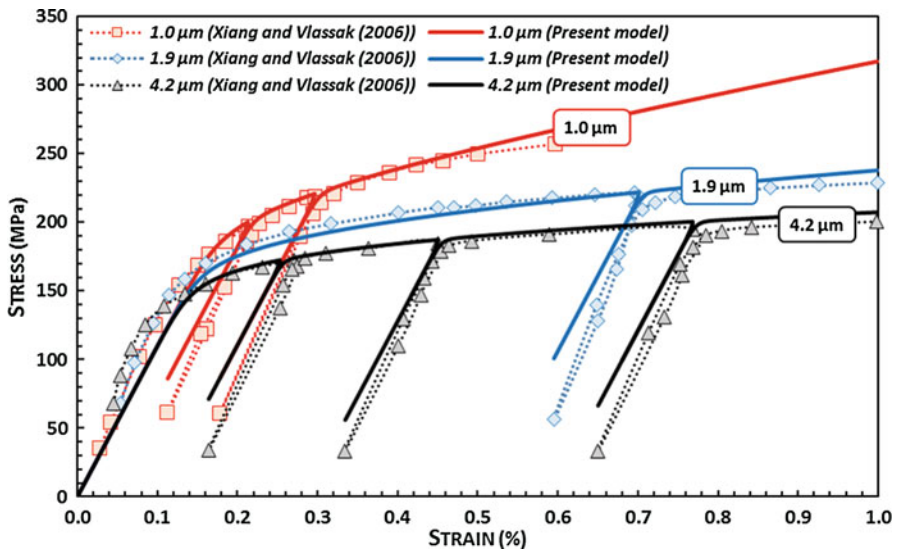


Fig. 4 The validation of the present strain gradient plasticity model by comparing to the experimental measurements from Xiang and Vlassak (2006) on the stress-strain response of the electroplated Cu thin films with the passivated layers on both sides (Reprinted with permission from Voyiadjis and Song 2017)

Image Correlation (DIC) system, for evaluating the mechanical properties of the Ni thin films at high temperatures. Dogbone-shaped specimens used in their experiments were made by Micro-electro mechanical system (MEMS) processes and the primary dimensions of the specimen are shown in Fig. 5.

The calibrated material parameters as well as the general parameters for Ni are presented in Table 1. The results of the microscale tensile tests at four different temperatures, i.e., 25 °C, 75 °C, 145 °C, and 218 °C, and corresponding numerical results from the present model (Voyiadjis and Song 2017) are shown in Fig. 6. As shown in this figure, it is clear from both the experimental and numerical results that

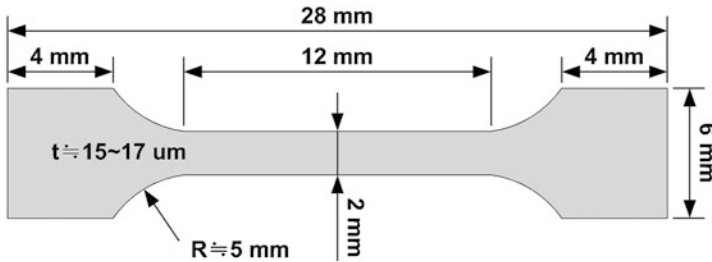


Fig. 5 The specimen dimensions for the experimental validation (Reprinted with permission from Voyiadjis and Song 2017)

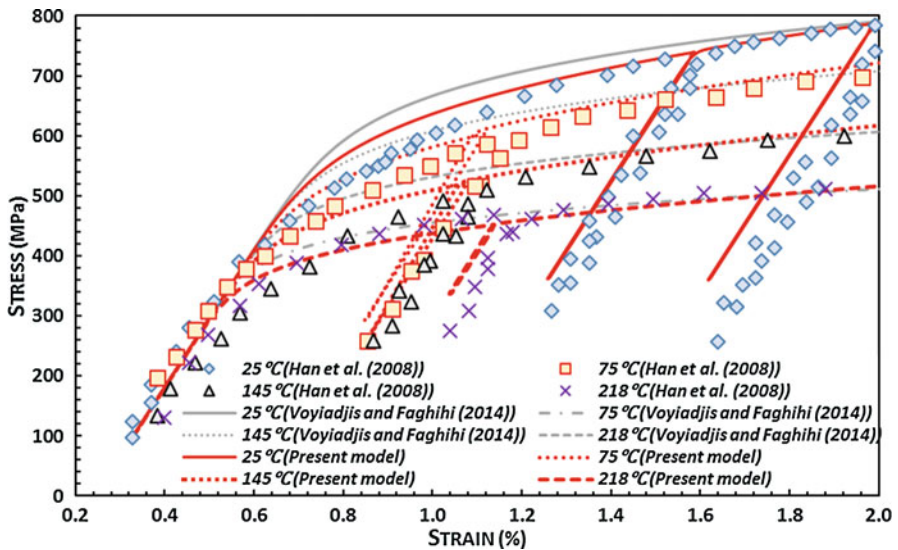


Fig. 6 The validation of the present strain gradient plasticity model (Voyiadjis and Song 2017) by comparing the numerical results with the model by Voyiadjis and Faghihi (2014) and the experimental measurements from Han et al. (2008) on the stress-strain response of Ni thin films (Reprinted with permission from Voyiadjis and Song 2017)

the Young's modulus is not affected by variations in temperature, while the yield and tensile strength decrease as the specimen temperature increases. In addition, Fig. 6 clearly shows that the Bauschinger effect is not affected very much by variations in the specimen temperature. Meanwhile, the calculated results of the present model (Voyiadjis and Song 2017) compare better to the experimental data than those of Voyiadjis and Faghihi (2014) (Fig. 6).

Stretch-Passivation Problem

The numerical solutions for the stretch-passivation problem with the two cases of the SGP model are presented in this section. The frameworks presented in this chapter represent the nonlocal flow rules in the form of partial differential equations when the microscopic force balances are integrated with the thermodynamically consistent constitutive equations. To interpret and analyze the physical phenomena characterized by the current frameworks is very complicated; in this sense, a one-dimensional numerical solution is presented first and extended to the two-dimensional one later in this chapter.

An initially uniform single grain with the size of L is used with two grain boundaries as shown in Fig. 7. The grain is assumed to be infinitely long along the x -direction and initially homogeneous; therefore, the solution depends only on the x -direction. In the one-dimensional stretch-passivation problem, the grain is deformed into the plastic regime by uniaxial tensile stretch with no constraint on plastic flow at the grain boundaries. At a certain point, the plastic flow is then constrained by blocking off the dislocations from passing out of the grain boundary, which leads the further plastic strain not to occur at the grain boundary.

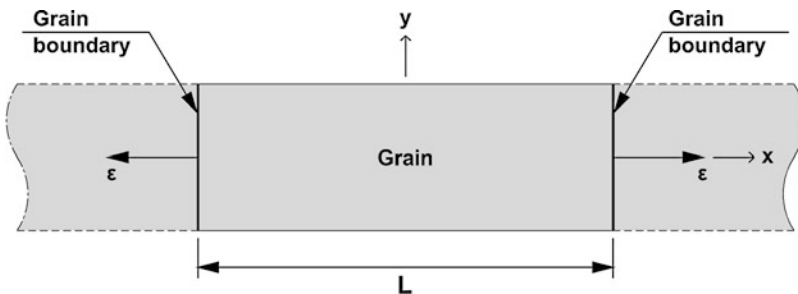


Fig. 7 One-dimensional model for a single grain with two grain boundaries (Reprinted with permission from Voyiadjis et al. 2017)

Numerical Results

In this section, an extensive numerical work is carried out based on the validated code in order to compare the results from the aforementioned two cases of the SGP model and to analyze the characteristics of the stress jump phenomenon. This section is largely composed of four subparts. In the first part, the occurrence of the stress jump phenomenon under the stretch-surface passivation condition is introduced in conjunction with three experiments used in section “[Experimental Validation of the Strain Gradient Plasticity Model.](#)” The second part is focused on indicating that the results are contrary to each other for the two cases of the SGP model. An extensive parametric study is also conducted in terms of the various material parameters, and the evolution of the free energy involving the stored energy of cold work and the dissipation potentials during the plastic deformation are discussed in the third part. In the last part, the two-dimensional simulation is also given to examine the gradient and grain boundary effect, the mesh sensitivity of the two-dimensional model, and the stress jump phenomenon.

The numerical results on the stress-strain behaviors of Al and Cu thin films for the SGP model with the corresponding dissipative microstress quantities are presented in Figs. 8 and 9, respectively. In both simulations, a significant stress jump

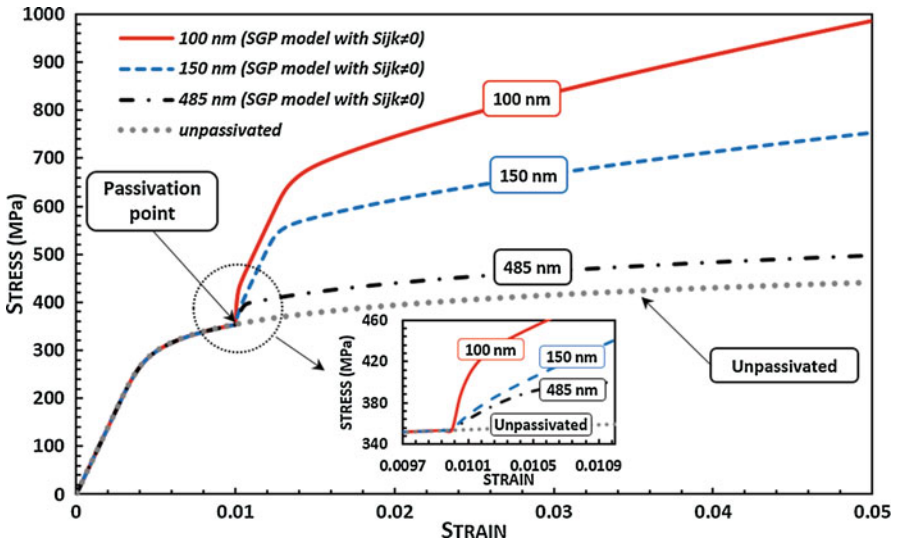


Fig. 8 A finite stress jump due to infinitesimal changes in the plastic strain. The numerical implementation of the SGP model with the dissipative potential dependent on $\dot{\epsilon}_{ij,k}^p$, i.e., $S_{ijk}^{\text{dis}} \neq 0$, is carried out based on the experiments of Haque and Saif (2003) with the various thicknesses of the Al thin films of 100 nm, 150 nm, and 485 nm (Reprinted with permission from Voyiadjis and Song 2017)

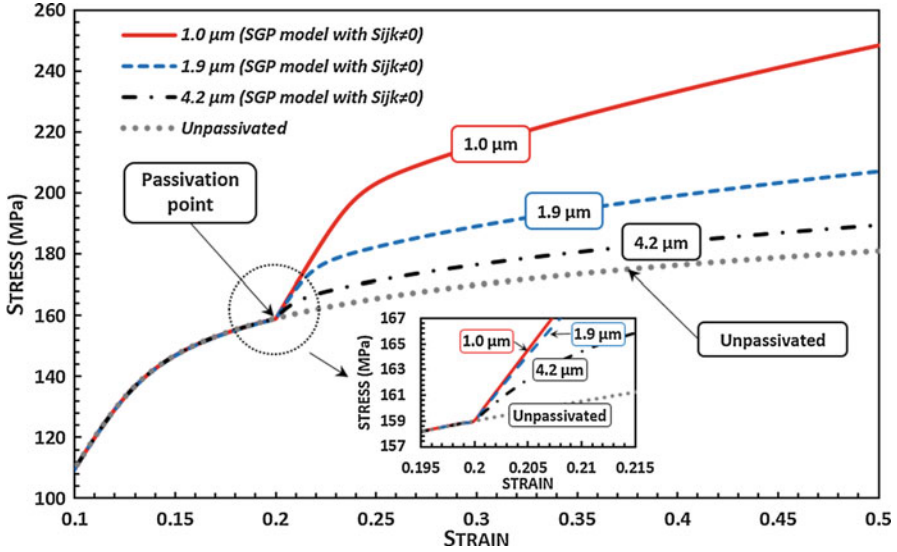


Fig. 9 A finite stress jump due to infinitesimal changes in the plastic strain. The numerical implementation of the SGP model with the dissipative potential dependent on $\dot{\epsilon}_{ij,k}^p$, i.e., $S_{ijk}^{dis} \neq 0$, is carried out based on the experiments of Xiang and Vlassak (2006) with the various thicknesses of the Cu thin films of 1.0 μm , 1.9 μm , and 4.2 μm (Reprinted with permission from Voyiadjis and Song 2017)

is observed at the onset of passivation. In particular, it is shown that the very first slopes immediately after the passivation increase as the film thicknesses decrease, viz. the dissipative length scales increase, in both simulations. Thus, the stress jump phenomenon is revealed to be highly correlated with the dissipative higher order microstress quantities S_{ijk}^{dis} .

Figure 10 shows the numerical results on the stress-strain behavior of Ni thin films for the SGP model with the dissipative microstress quantities. As shown in this figure, the magnitudes of the stress jump are less than expected in all cases since the dissipative length scale ℓ_{dis} is set 0.1 which is much smaller than the energetic length scale $\ell_{en} = 1.0$. Nevertheless, the very first slopes immediately after the passivation are calculated as $E_{25^\circ\text{C}} = 58.0$ GPa, $E_{75^\circ\text{C}} = 59.2$ GPa, $E_{145^\circ\text{C}} = 72.6$ GPa, and $E_{218^\circ\text{C}} = 105.0$ GPa, respectively, and this shows the responses immediately after the passivation gets gradually closer to the elastic response $E = 115$ GPa as the temperature increases.

The numerical implementations to specify whether or not the stress jump phenomenon occurs under the stretch-surface passivation have been hitherto conducted within the framework of the SGP model with $S_{ijk}^{dis} \neq 0$. Hereafter, the numerical simulations are given more focus on the direct comparison of the material response on the stress-strain curves between the two cases of the SGP models. The material parameters used for these implementations are presented in Table 2.

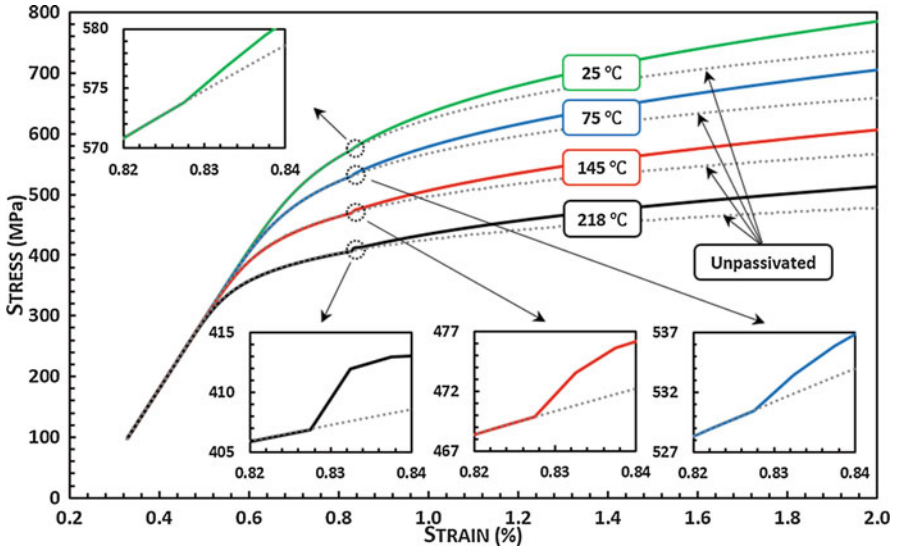


Fig. 10 A finite stress jump due to infinitesimal changes in the plastic strain. The numerical implementation of the SGP model with the dissipative potential dependent on $\dot{\epsilon}_{ij,k}^p$, i.e., $S_{ijk}^{dis} \neq 0$, is carried out based on the experiments of Han et al. (2008) at the various temperatures of the Ni thin films of 25 °C, 75 °C, 145 °C, and 218 °C. The circles on the curve indicate the passivation point (Reprinted with permission from Voyiadjis and Song 2017)

Table 2 Material parameters for the numerical simulation (Reprinted with permission from Voyiadjis and Song (2017))

For grain (bulk)		
E	Elastic modulus for isotropic linear elasticity	100 (GPa)
σ_y	Yield stress	100 (MPa)
h	Hardening material parameter	200 (MPa)
$\dot{\gamma}_0$	Reference effective plastic strain rate	0.04
r	Nonlinear hardening material parameter	0.1
m	Non-negative rate sensitivity parameter	0.3
T_0	Initial temperature	77 (K)
T_y	Thermal material parameter	1,000 (K)
n	Temperature sensitivity parameter	0.6
ρ	Density	0.8570 g • cm ⁻³
c_ϵ	Specific heat capacity at constant stress	0.265 (J/g K)
For grain boundary (interface)		
σ_y^I	Interfacial Yield stress	150 (MPa)
h^I	Interfacial hardening material parameter	300 (MPa)
$\dot{\gamma}_0^I$	Interfacial reference effective plastic strain rate	0.04
n^I	Interfacial temperature sensitivity parameter	0.1
m^I	Interfacial rate sensitivity parameter	1.0
T_y^I	Interfacial thermal material parameter	700 (K)

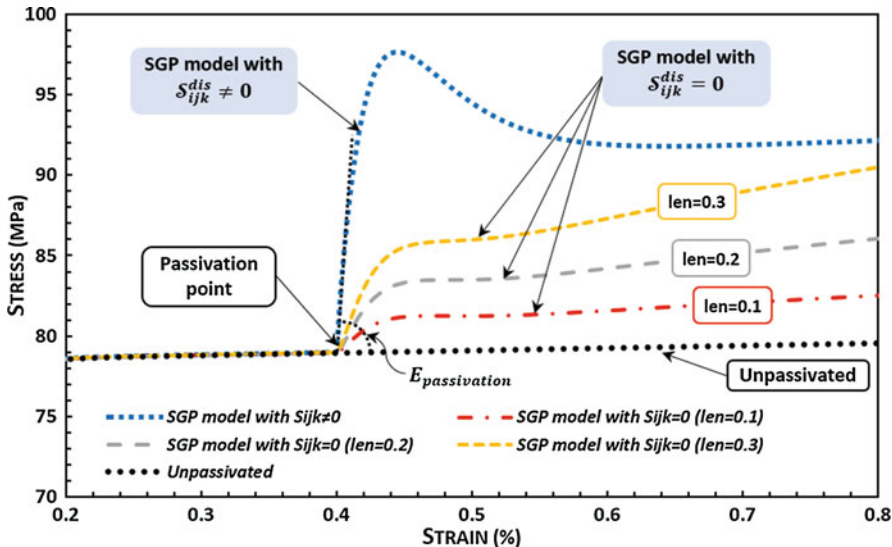


Fig. 11 Comparison of the results from two different cases of the present SGP model with the dissipation potential dependent on $\dot{\epsilon}_{ij,k}^p$, i.e., $S_{ijk}^{dis} \neq 0$, and with the dissipative potential independent on $\dot{\epsilon}_{ij,k}^p$, i.e., $S_{ijk}^{dis} = 0$. The results for the latter case are computed with three different values of energetic length scales $\ell_{en} = 0.1, 0.2$ and 0.3 (Reprinted with permission from Voyiadjis and Song 2017)

Figure 11 clearly shows this point by comparing the results from the two cases. The behavior of the SGP model with $S_{ijk}^{dis} \neq 0$ is in stark contrast with that of the SGP case with $S_{ijk}^{dis} = 0$ after the passivation point. A significant stress jump with the slope $E_{passivation}$ similar to the modulus of elasticity E is shown in the SGP case with $S_{ijk}^{dis} \neq 0$. On the other hand, no elastic stress jump is observed in the SGP case with $S_{ijk}^{dis} = 0$. This result is exactly in agreement with the prediction in Fleck et al. (2014, 2015). In the case that the dissipative potential is independent of $\dot{\epsilon}_{ij,k}^p$, the contribution from the plastic strain gradients is entirely energetic as can be seen in section “Dissipation Potential and Dissipative Thermodynamic Microstresses.” Both the increase in the yield strength in the early stages of passivation and subsequent hardening due to the effects of the plastic strain gradient are observed along with the increase of the energetic length scale as shown in Fig. 11.

The comparison of the results from the case of the present SGP model by Voyiadjis and Song (2017) with the dissipative potential dependent on $\dot{\epsilon}_{ij,k}^p$ is shown in Fig. 12a with various dissipative length scales, i.e., $\ell_{dis} = 0.1, 0.3, 0.5, 1.0, 1.5,$ and 2.0 . As can be seen in this figure, the magnitude of the stress jump significantly increases as the dissipative length scale increases, on the other hand, the stress jump phenomenon disappears as the dissipative length scale tends to zero. This is because the dissipative higher order microstress quantities S_{ijk}^{dis} , which is the main cause of the stress jump, vanishes when the dissipative length scale ℓ_{dis} is

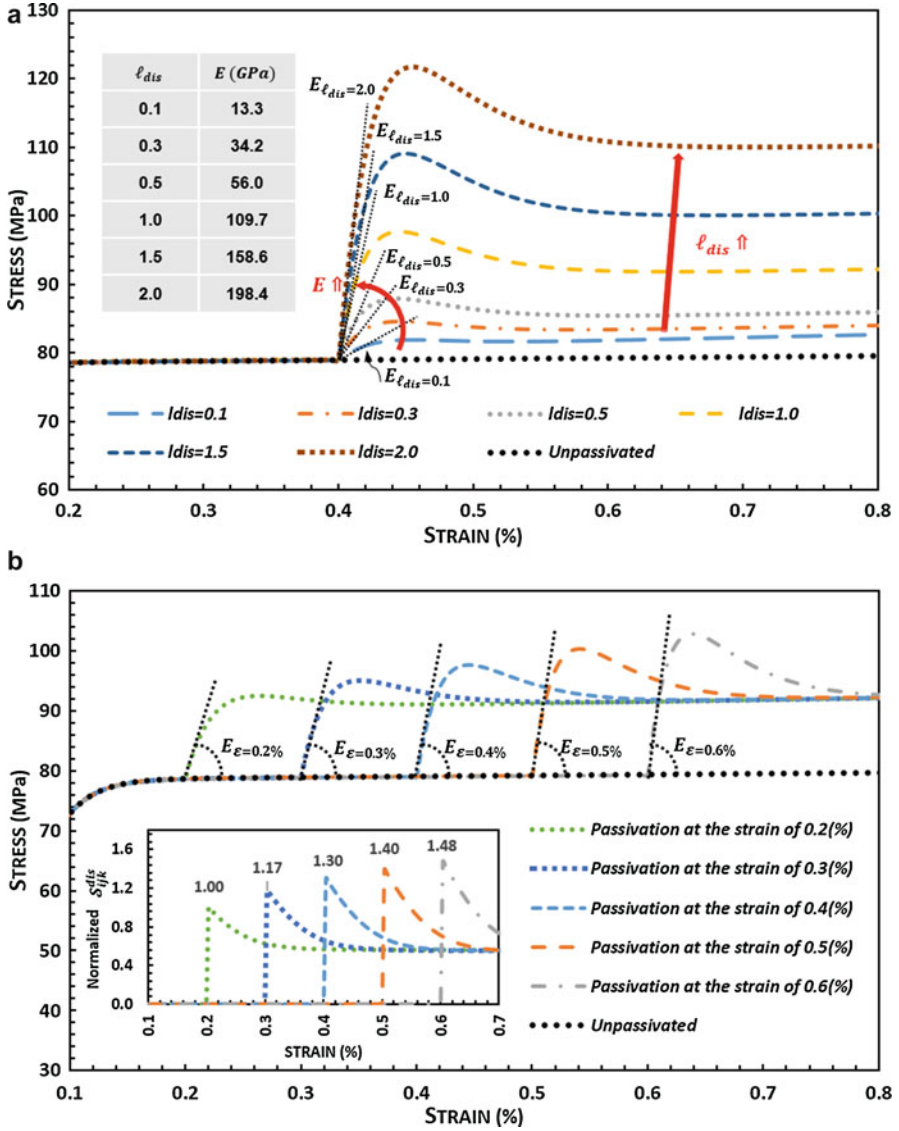


Fig. 12 Comparison of the results from the case of the present SGP model with the dissipative potential dependent on $\dot{\epsilon}_{ij,k}^p$, i.e., $S_{ijk}^{dis} \neq 0$: (a) with various dissipative length scales ($\ell_{dis} = 0.1, 0.3, 0.5, 1.0, 1.5, \text{ and } 2.0$) and (b) for various passivation points ($\epsilon = 20.2\%, 0.3\%, 0.4\%, 0.5\%, \text{ and } 0.6\%$ with identical energetic and dissipative length scales) (Reprinted with permission from Voyiadjis and Song 2017)

equal to zero. It is worth noticing that the very first behavior immediately after the passivation indicates a substantial difference with varying dissipative length scales. As the dissipative length scales increase from 0.1 to 2.0, the corresponding slopes of the very first response E also increase from 13.3 to 198.4 GPa. The main reason of this phenomenon in terms of the dissipative length scale is also because the dissipative higher order microstress quantities $\mathcal{S}_{ijk}^{\text{dis}}$ sharply increases with increasing ℓ_{dis} .

The comparison of the results from the case of the SGP model with the dissipative potential dependent on $\varepsilon_{ij,k}^p$ is shown in Fig. 12b for various passivation points. The energetic and dissipative length scales are set identical in all cases. The magnitudes of the stress jump with the magnitudes of $\varepsilon = 20.2\%$, 0.3%, 0.4%, 0.5%, and 0.6% are obtained as 13.8 MPa, 16.2 MPa, 18.6 MPa, 21.1 MPa, and 23.5 MPa, respectively, and the values normalized by the value of the magnitude $\varepsilon = 20.2\%$ are calculated as 1.00, 1.17, 1.35, 1.52, and 1.70, respectively. The normalized higher order microstress quantities $\mathcal{S}_{ijk}^{\text{dis}}$ with the magnitudes of $\varepsilon = 20.2\%$, 0.3%, 0.4%, 0.5%, and 0.6% are obtained as 1.00, 1.17, 1.30, 1.40, and 1.48, respectively, as shown in Fig. 12b. Thus, it is worth noticing that the stress jump phenomenon is highly correlated with the dissipative higher order microstress quantities $\mathcal{S}_{ijk}^{\text{dis}}$. In addition, the very first responses immediately after the passivation also make a substantial difference with varying passivation points. The slopes of the very first responses for the magnitudes of $\varepsilon = 0.2\%$, 0.3%, 0.4%, 0.5%, and 0.6% are calculated as 64.8 GPa, 87.1 GPa, 109.7 GPa, 132.5 GPa, and 155.5 GPa, respectively.

The effects of various parameters on the mechanical behavior of the stretch-surface passivation problem are investigated by using a one-dimensional finite element. The numerical results reported in this parametric study are obtained by using the values of the material parameters in Table 2 unless it is differently mentioned.

The stress-strain graphs for various values of the hardening material parameter h are shown in Fig. 13a with two different cases of the SGP model. For the case with $\mathcal{S}_{ijk}^{\text{dis}} \neq 0$, the slopes of the very first response immediately after the passivation are obtained as $E_h = 109.7$ GPa in all simulations, and the corresponding magnitudes of the stress jump for each simulation are, respectively, obtained as $\sigma'_{h=100\text{MPa}} = 18.6$ MPa, $\sigma'_{h=200\text{MPa}} = 18.6$ MPa, $\sigma'_{h=300\text{MPa}} = 18.7$ MPa, $\sigma'_{h=400\text{MPa}} = 18.7$ MPa, and $\sigma'_{h=500\text{MPa}} = 18.7$ MPa. There is little difference between all the simulations. For the case with $\mathcal{S}_{ijk}^{\text{dis}} = 0$, no stress jump phenomena are observed in all simulations.

The effects of the nonnegative rate sensitivity parameter m on the stress-strain behavior for the two cases of the SGP model are represented in Fig. 13b. It is clearly shown in this figure that by increasing the rate sensitivity parameter, the stress jump phenomena are significantly manifested in the case of the SGP model with $\mathcal{S}_{ijk}^{\text{dis}} \neq 0$ in terms of both the slope of the very first response immediately after the passivation and the corresponding magnitude of the stress jump. In the case of the SGP model with $\mathcal{S}_{ijk}^{\text{dis}} = 0$, on the other hand, the material behavior is not affected a lot by the rate sensitivity parameter m . In the SGP model with $\mathcal{S}_{ijk}^{\text{dis}} \neq 0$, the slopes of

the very first response immediately after the passivation (E_m) and the corresponding magnitudes of the stress jump (σ'_m) are obtained as $E_{m=0.2} = 96.7$ GPa, $E_{m=0.25} = 103.0$ GPa, $E_{m=0.3} = 109.7$ GPa, $E_{m=0.35} = 116.5$ GPa and $\sigma'_{m=0.2} = 14.8$ MPa, $\sigma'_{m=0.25} = 16.7$ MPa, $\sigma'_{m=0.3} = 18.6$ MPa, $\sigma'_{m=0.35} = 20.6$ MPa, respectively. Thus, it is clearly shown that both the slope of the very first response immediately after the passivation and the corresponding magnitude of the stress jump increase as the nonnegative rate sensitivity parameter m increases.

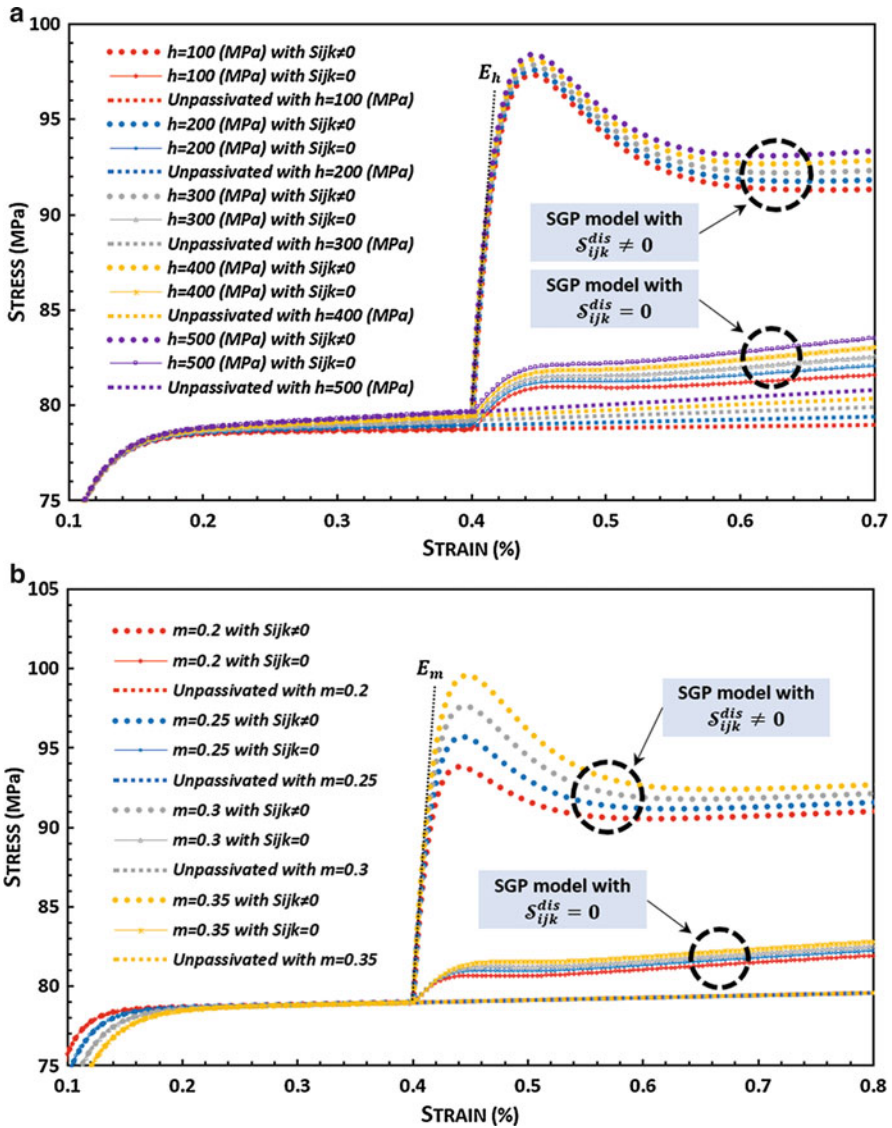


Fig. 13 (continued)

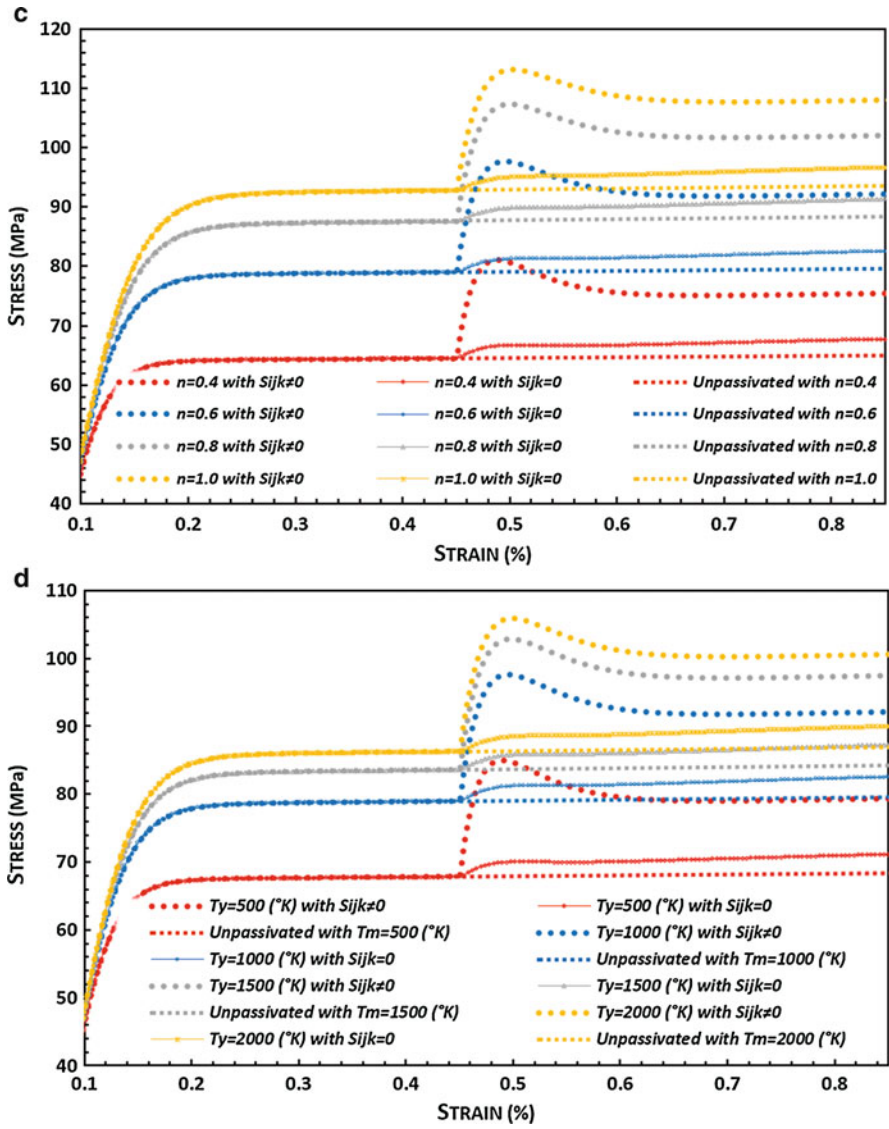


Fig. 13 (continued)

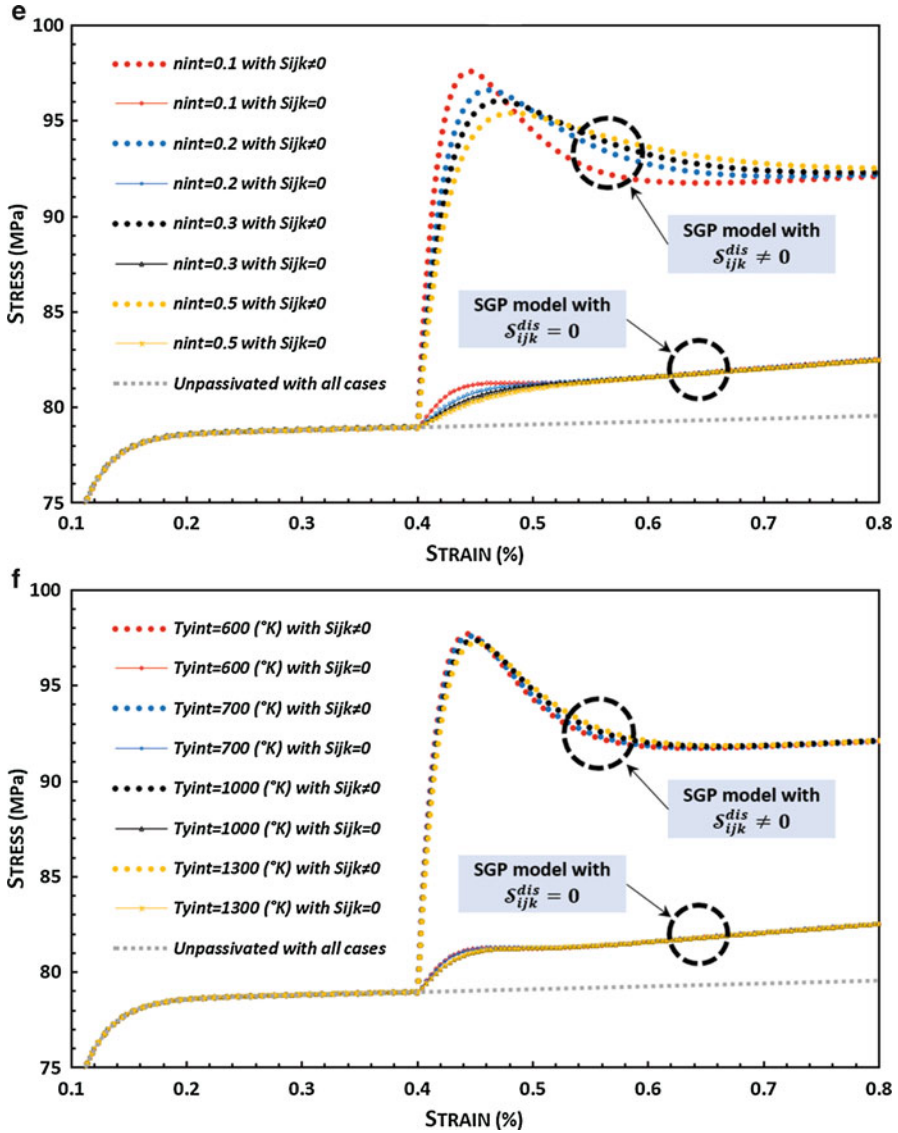


Fig. 13 Comparison of the results from the case of the present SGP model with the dissipative potential dependent on $\dot{\epsilon}_{ijk}^p$, i.e., $S_{ijk}^{dis} \neq 0$ with the effects of: **a** the hardening material parameter h (100 MPa, 200 MPa, 300 MPa, 400 MPa, and 500 MPa), **b** the nonnegative rate sensitivity parameter m (0.2, 0.25, 0.3, and 0.35), **c** the temperature sensitivity parameter n (0.4, 0.6, 0.8, and 1.0), **d** the thermal material parameter T_y ($500^{\circ}K$, $1,000^{\circ}K$, $1,500^{\circ}K$, and $2,000^{\circ}K$), **e** the interfacial temperature sensitivity parameter n^I (0.1, 0.2, 0.3 and 0.5), **f** the interfacial thermal material parameter T_y^I ($600^{\circ}K$, $700^{\circ}K$, $1,000^{\circ}K$, and $11,300^{\circ}K$) (Reprinted with permission from Voyiadjis and Song 2017)

The effects of the temperature sensitivity parameter n on the stress-strain behavior for the two cases of the SGP model are represented in Fig. 13c. It is clearly shown in this figure that the yield stress significantly increases as the temperature sensitivity parameter n increases, while the strain hardening is not influenced a lot by this parameter. This is because the temperature affects the strain hardening mechanism through the dislocation forest barriers, while the backstress, i.e., energetic gradient hardening, is almost independent of the temperature. Meanwhile, the temperature sensitivity parameter n significantly affects the stress-strain response in the case of the SGP model with $\mathcal{S}_{ijk}^{\text{dis}} \neq 0$. In this case, the slopes of the very first response immediately after the passivation are obtained as $E_{n=0.4} = 112.5$ GPa, $E_{n=0.6} = 109.7$ GPa, $E_{n=0.8} = 107.8$ GPa, and $E_{n=1.0} = 106.6$ GPa, and the corresponding magnitudes of the stress jump are obtained as $\sigma'_{n=0.4} = 16.6$ MPa, $\sigma'_{n=0.6} = 18.6$ MPa, $\sigma'_{n=0.8} = 19.7$ MPa, and $\sigma'_{n=1.0} = 20.4$ MPa, respectively. Thus, in contrast with the rate sensitivity parameter m , the slope of the very first response immediately after the passivation decreases, while the corresponding magnitude of the stress jump increases as the temperature sensitivity parameter n increases.

The effects of the thermal material parameter T_y on the stress-strain behavior with the two cases of the SGP model are represented in Fig. 13d. Similar to the temperature sensitivity parameter n , the yield stress significantly increases as the thermal material parameter T_y increases, while the strain hardening is not influenced a lot by this parameter. In the case of the SGP model with $\mathcal{S}_{ijk}^{\text{dis}} \neq 0$, the slopes of the very first response immediately after the passivation (E_{T_y}) and the corresponding magnitudes of the stress jump (σ'_{T_y}) are obtained as $E_{T_y=500^\circ K} = 111.9$ GPa, $E_{T_y=1000^\circ K} = 109.7$ GPa, $E_{T_y=1500^\circ K} = 108.7$ GPa, $E_{T_y=2000^\circ K} = 108.1$ GPa and $\sigma'_{T_y=500^\circ K} = 17.1$ MPa, $\sigma'_{T_y=1000^\circ K} = 18.6$ MPa, $\sigma'_{T_y=1500^\circ K} = 19.2$ MPa, $\sigma'_{T_y=2000^\circ K} = 19.6$ MPa, respectively. These results from the simulations with various thermal material parameter T_y are very similar to those with the temperature sensitivity parameter n , in the sense that the slope of the very first response immediately after the passivation decreases, while the corresponding magnitude of the stress jump increases as the temperature sensitivity parameter n increases.

The effects of the interfacial temperature sensitivity parameter n^I on the stress-strain behavior for the two cases of the SGP model are presented in Fig. 13e. It is clearly shown in this figure that increasing interfacial temperature sensitivity parameter makes the grain boundary (interface) harder and results in less variation of the stress jump in both cases of the SGP model with $\mathcal{S}_{ijk}^{\text{dis}} \neq 0$ and $\mathcal{S}_{ijk}^{\text{dis}} = 0$. In the case of the SGP model with $\mathcal{S}_{ijk}^{\text{dis}} \neq 0$, the slopes of the very first response immediately after the passivation are obtained as $E_{n^I=0.1} = 109.7$ GPa, $E_{n^I=0.2} = 80.4$ GPa, $E_{n^I=0.3} = 70.3$ GPa, and $E_{n^I=0.5} = 62.1$ GPa, and the corresponding magnitudes of the stress jump are obtained as $\sigma'_{n^I=0.1} = 18.6$ MPa, $\sigma'_{n^I=0.2} = 17.6$ MPa, $\sigma'_{n^I=0.3} = 17.1$ MPa, and $\sigma'_{n^I=0.5} = 16.4$ MPa, respectively. From these results, it is easily observed that the slope of the very first response after the passivation and the corresponding magnitude of the stress jump increase as the interfacial temperature sensitivity parameter n^I decreases. In addition, the variations along with the different cases are shown to be more drastic by decreasing the

interfacial temperature sensitivity parameter n^I . Similarly, for the case of the SGP model with $\mathcal{S}_{ijk} = 0$, decreasing the interfacial temperature sensitivity parameter makes the variation more radically as shown in Fig. 13e.

The effects of the interfacial thermal material parameter T_y^I on the stress-strain behavior for the two cases of the SGP model are represented in Fig. 13f. As can be seen in this figure, the overall characteristic from the simulation results with various T_y^I is similar to those with various n^I in the sense that increasing T_y^I makes the grain boundary (interface) harder and results in less variation of the stress jump in both cases of the SGP model with $\mathcal{S}_{ijk}^{\text{dis}} \neq 0$ and $\mathcal{S}_{ijk}^{\text{dis}} = 0$. In the case of the SGP model with $\mathcal{S}_{ijk}^{\text{dis}} \neq 0$, the slopes of the very first response immediately after the passivation are obtained as $E_{T_y^I=600^\circ K} = 114.0$ GPa, $E_{T_y^I=700^\circ K} = 109.7$ GPa, $E_{T_y^I=1000^\circ K} = 101.7$ GPa, and $E_{T_y^I=1300^\circ K} = 97.1$ GPa, and the corresponding magnitudes of the stress jump are obtained as $\sigma'_{T_y^I=600^\circ K} = 18.7$ MPa, $\sigma'_{T_y^I=700^\circ K} = 18.6$ MPa, $\sigma'_{T_y^I=1000^\circ K} = 18.4$ MPa, and $\sigma'_{T_y^I=1300^\circ K} = 18.3$ MPa, respectively.

Based on the calibrated model parameters of Ni (Table 1), the evolution of the various potentials during the plastic deformation are investigated with four different temperatures, i.e., 25 °C, 75 °C, 145 °C, and 218 °C, in this section. Figure 14 shows (a) the variation of the plastic strain dependent free energies (Ψ_1^d , $\Psi_1^{d,R}$ and $\Psi_1^{d,NR}$), (b) plastic strain dependent dissipation rate \mathcal{D}_1 (i.e., plastic strain dependent term in Eq. 58), (c) plastic strain gradient-dependent free energy (Ψ_2^d) and dissipation rate \mathcal{D}_2 (i.e., plastic strain gradient dependent term in Eq. 58), and (d) amount of stored energy ($\Psi_1^d + \Psi_2^d$) and dissipated energy ($\mathbb{D} = \int (\mathcal{D}_1 + \mathcal{D}_2) dt$).

As can be seen in Fig. 14a, both free energies Ψ_1^d and $\Psi_1^{d,NR}$ decrease as the temperature increases. The stored energy of cold work $\Psi_1^{d,R}$ is also presented in this figure. One can observe that Ψ_1^d and $\Psi_1^{d,NR}$ have a strong temperature dependency, while the stored energy of cold work has no variation with varying temperatures. In addition, the stored energy of cold work tends to saturate after some critical point since the stored energy of cold work is proportional to the dislocation density that remains constant after the aforementioned critical point. Meanwhile, it is shown in Fig. 14b,c that rates of dissipation, \mathcal{D}_1 and \mathcal{D}_2 , are dependent on the temperature such that both \mathcal{D}_1 and \mathcal{D}_2 increase with decreasing temperatures while Ψ_2^d shows no temperature dependency. The amount of stored and dissipated energies is shown in Fig. 14d. As it is shown in this figure, the amount of stored energy is larger than the dissipated energy and both decrease as the temperature increases.

Lastly, the one-dimensional finite element implementation is extended to the two-dimensional one. The simple tension problem of the square plate is solved to study the strain gradient effects, the mesh sensitivity of the model with the three cases according to the number of elements (100, 400, and 1,600 elements), and the stress jump phenomena under the abrupt surface passivation. Each edge of the plate has a length of w and the material parameters in Table 2 are used again for these simulations.

Figure 15a, b shows the stress-strain behavior of the plate with the various energetic length scales ($\ell_{en}/w = 0.0, 0.1, 0.3, 0.5, 0.7,$ and 1.0 with $\ell_{dis}/w = 0.0$) and dissipative length scales ($\ell_{dis}/w = 0.0, 0.1, 0.2, 0.3, 0.4,$

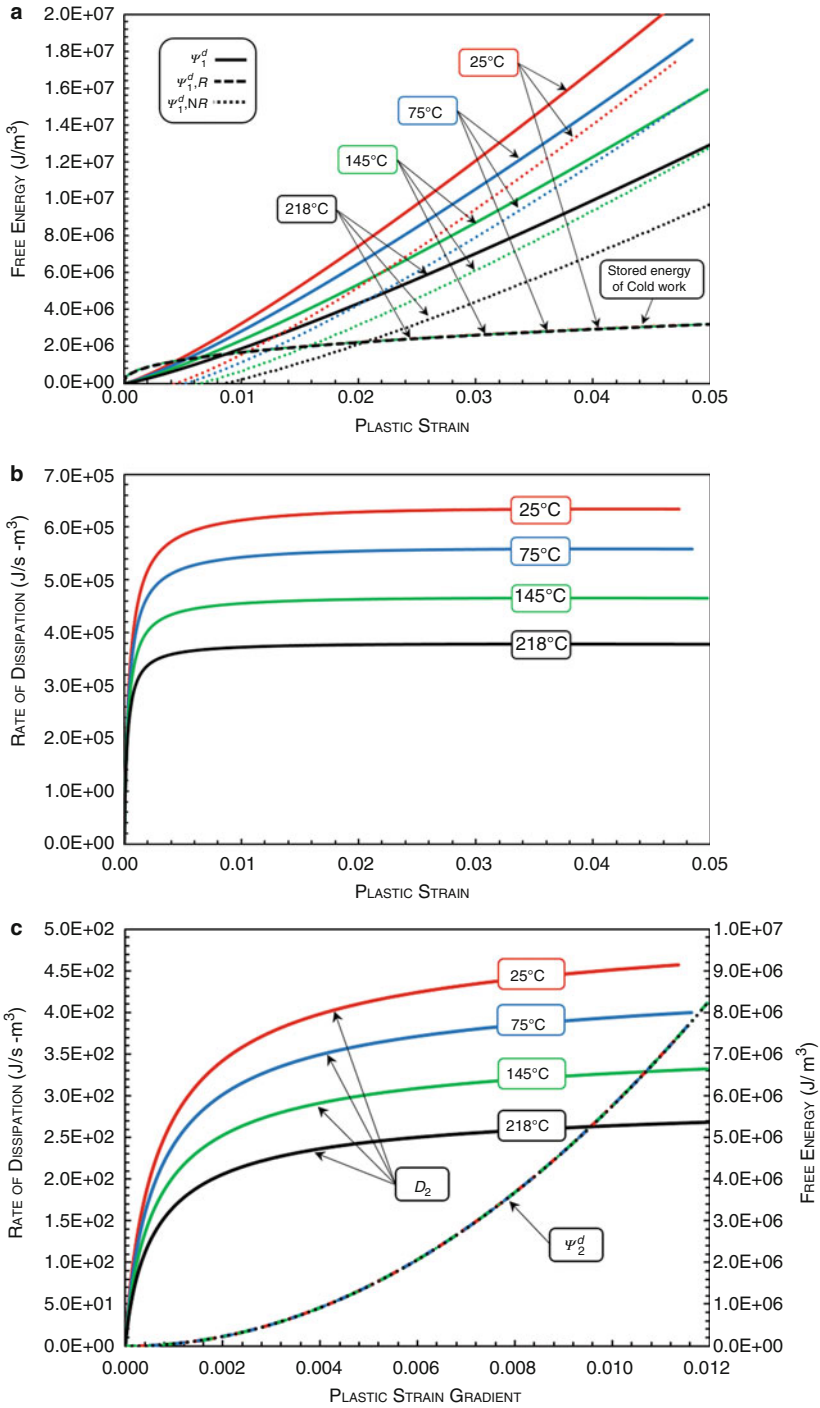


Fig. 14 (continued)

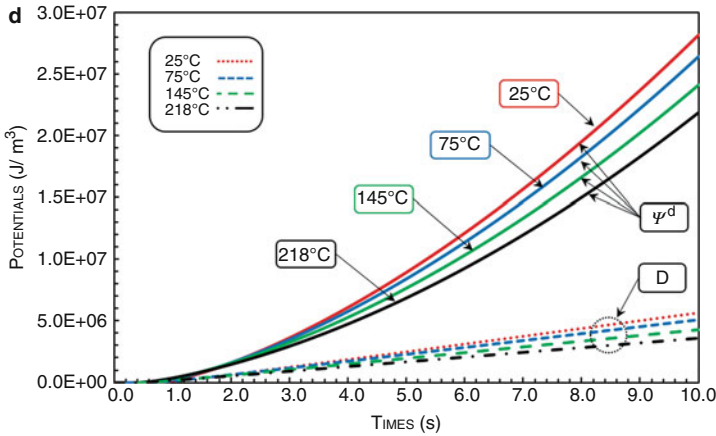


Fig. 14 The evolution of free energy and dissipation potentials based on the calibrated model parameters (Han et al. 2008): (a) plastic strain-dependent free energy, (b) plastic strain-dependent dissipation rate, (c) plastic strain gradient-dependent free energy and dissipation rate (The primary y-axis on the LHS: \mathcal{D}_2 , the secondary y-axis on the RHS: Ψ_2^d), and (d) amount of stored and dissipated energies (Reprinted with permission from Voyiadjis and Song 2017)

and 0.5 with $\ell_{en}/w = 0.0$), respectively. In common with the one-dimensional simulations, the numerical results show the energetic hardening as the energetic length scale increases as well as the dissipative strengthening as the dissipative length scale increases.

Figure 15c shows the grain boundary effect of the square plate. It is well known that the grain boundary blocks the dislocation movement, which, in turn, leads to the strengthening of the material. The energetic and dissipative length scales reported in these simulations are set zero and the microclamped condition is imposed at the grain boundary, which is indicated by the bold line in the figure. The strengthening caused by increasing the grain boundary area is well observed as expected. For the simulations presented in Fig. 15a–c, 20×20 elements are used.

The mesh sensitivity of the two-dimensional numerical model is examined in terms of the energetic and dissipative length scales with 10×10 , 20×20 and 40×40 mesh elements in Fig. 15d, e, respectively. Figure 15d shows the stress-strain behavior of the plate with the various energetic length scales ($\ell_{en}/w = 0.1, 0.5$ and 1.0 with $\ell_{dis}/w = 0$) compared to those in the absence of the gradient effects ($\ell_{en}/w = \ell_{dis}/w = 0$). The numerical results without the gradient terms significantly show the mesh sensitivity as expected, while for all nonzero values of ℓ_{en}/w , the numerical solutions show the mesh-independent behavior. In addition, Fig. 15d also shows the energetic hardening as the energetic length scale increases. The mesh-independent behavior is also observed with varying dissipative length scales $\ell_{dis}/w = 0.1, 0.2$ and 0.3 with $\ell_{en}/w = 0$ in Fig. 15e as with the case for the energetic length scale. The dissipative strengthening is also observed in this figure.

The stress jump phenomenon hitherto extensively studied in the one-dimensional finite element implementation is also examined for the two-dimensional simulation;

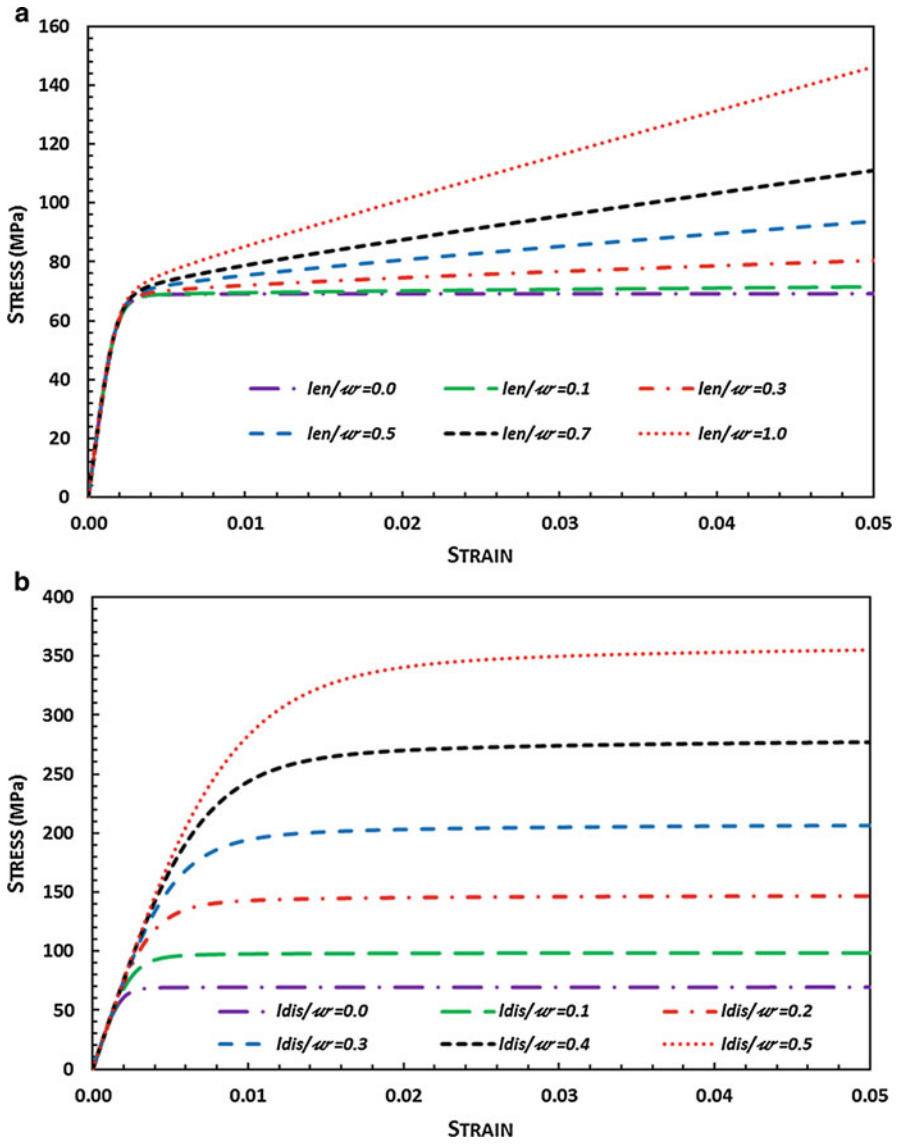


Fig. 15 (continued)

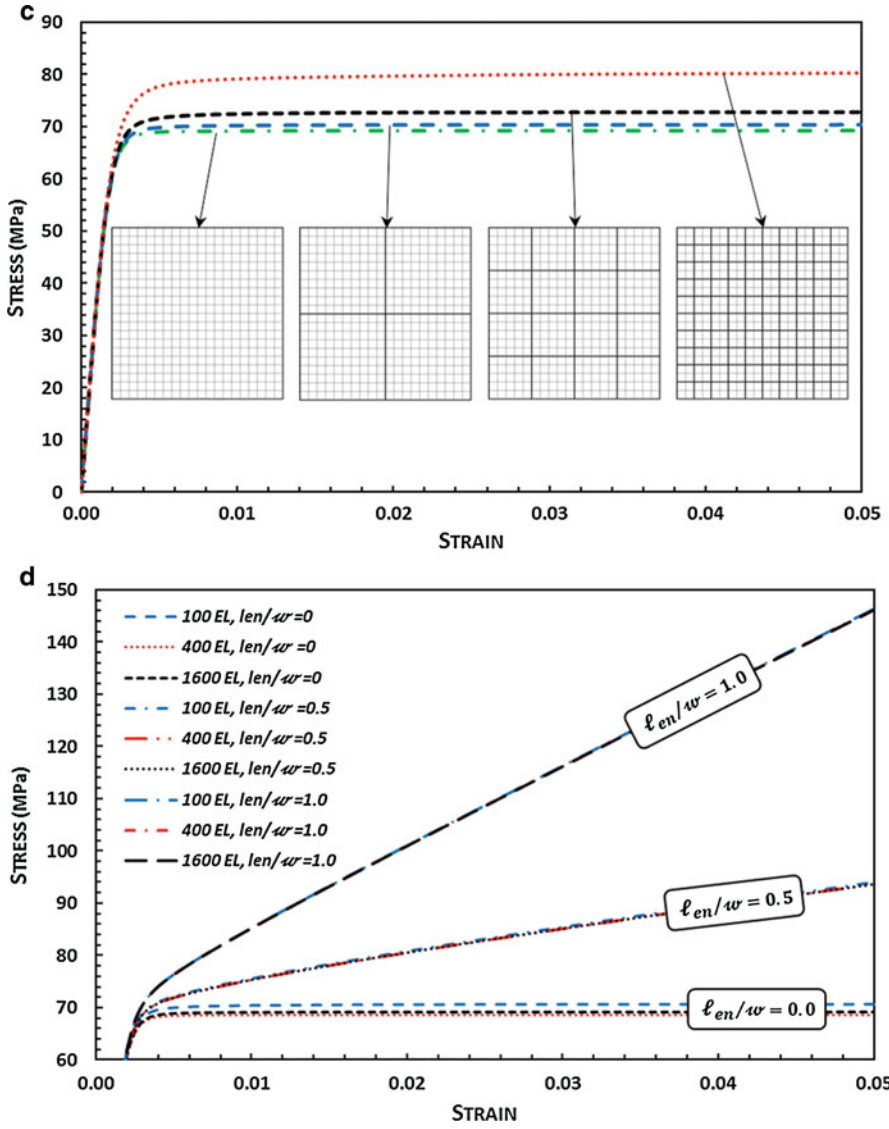


Fig. 15 (continued)

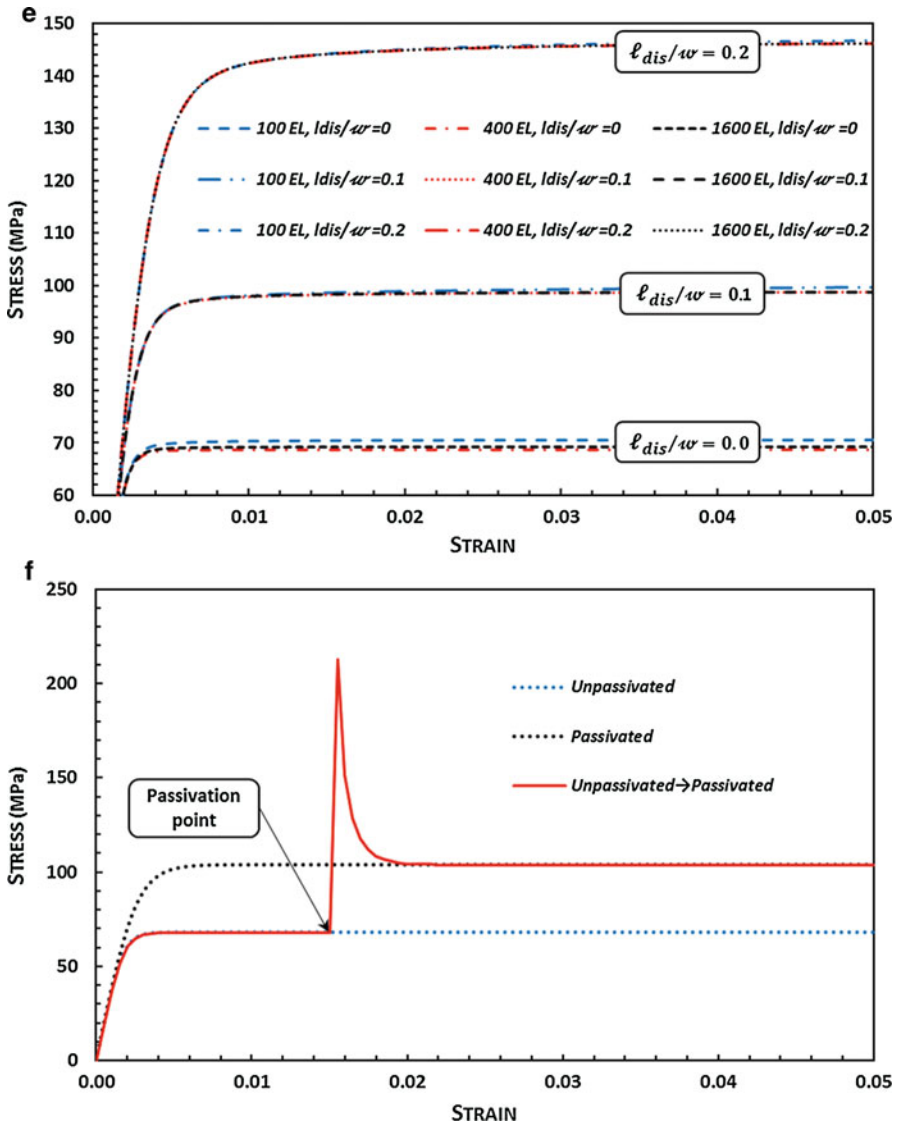


Fig. 15 The stress-strain behavior of the square plate: (a) the energetic hardening, (b) the dissipative strengthening, (c) the grain boundary effect, (d) the mesh sensitivity with varying l_{en} , (e) the mesh sensitivity with varying l_{dis} , and (f) the stress jump phenomenon (Fig. 15c, f: Reprinted with permission from Voyiadjis and Song (2017))

20×20 elements with $\ell_{\text{en}}/w = 0.0$ and $\ell_{\text{dis}}/w = 0.2$ are used for this simulation. Figure 15f shows the material behaviors of the unpassivated plate, the passivated plate, and the plate abruptly passivated at some point. As can be seen in this figure, the stress jump is significantly observed by the numerical results, which are identical to the one-dimensional case.

Conclusions

In this chapter, a phenomenological thermodynamic-based higher order gradient plasticity theory is presented and applied to the stretch-surface passivation problem for investigating the material behavior under the nonproportional loading condition. The thermodynamic potentials such as the Helmholtz free energy and the dissipation potential are established based on the concepts of the dislocation interaction mechanism and the thermal activation energy. The microstructural interface effect between two grains is also incorporated into the formulation, such that the present interfacial flow rule is able to account for the energy storage at the interface caused by the dislocation pile up as well as the energy dissipation through the interface caused by the dislocation transfer. The formulation is tested for two cases in the presence of the dissipative higher order microstress quantities S_{ijk}^{dis} . In the first case, the dissipation potential is dependent on the gradients of the plastic strain rate $\dot{\epsilon}_{ij,k}^p$; as a result, S_{ijk}^{dis} does not have a value of zero in this formulation. In the second case the dissipation potential is independent of $\dot{\epsilon}_{ij,k}^p$, which in turn, S_{ijk}^{dis} does not exist. It is noticed by Fleck et al. (2014, 2015) that S_{ijk}^{dis} always gives rise to the stress jump phenomenon, which causes a controversial dispute in the field of SGP theory with respect to whether it is physically acceptable or not, under the nonproportional loading condition.

Prior to exploring the effect of the dissipative higher order microstress quantities S_{ijk}^{dis} on the stress-strain behavior for the two cases of the SGP model with and without S_{ijk}^{dis} , the present model and corresponding finite element code by Voyiadjis and Song (2017) are validated by comparing with three sets of small-scale experiments. Particularly, each set of three experiments involving Al, Cu, and Ni are selected, respectively, to examine the applicability of the present finite element implementation to the various kinds of materials. The first experiment, which was performed by Haque and Saif (2003), is the uniaxial tensile testing of nanoscale freestanding Al thin films to explore the effect of strain gradient in 100 nm, 150 nm, and 485 nm thick specimens with average grain size of 50 nm, 65 nm, and 212 nm, respectively. The results clearly show the size effect on the stress-strain curves of the Al thin films; in addition, the calculated results of the present SGP model display a tendency to be more coincident to the experimental data than those of Voyiadjis and Faghihi (2014). For the second experimental validation, the experimental work of Xiang and Vlassak (2006) on the size effect in electroplated Cu thin films with various microscale thicknesses is selected since the effect of passivation on the stress-strain behavior of the Cu thin film is also considered in their work. The

stress-strain curves from the numerical results of the present SGP (Voyiadjis and Song 2017) model are in good agreement with the experimental measurements. The size effect according to the variation of the film thicknesses is also well observed from the results. For the third experimental validation, the microtensile test on the temperature effect on Ni thin films by Han et al. (2008) is employed since, in the nano- or microsystems, the effect of the higher order gradient on temperature also needs to be considered for the fast transient behavior. From both the experimental and numerical results, it is shown that the Young's modulus is not affected by the variations in temperature, while the yield and tensile strength decrease as the specimen temperature increases. The calculated results of the present model (Voyiadjis and Song 2017) compare better to the experimental data than those of Voyiadjis and Faghihi (2014).

The numerical simulation part is largely composed of four subparts. The main purpose of the first part is to examine the occurrence of the stress jump phenomenon under the stretch-surface passivation condition in conjunction with the aforementioned three experiments. In all simulations, a stress jump is clearly observed at the onset of passivation. The second part is carried out in order to clearly show the results to be contrary to each other from the two cases of the SGP model. The conclusion in this part is drawn such that a significant stress jump with the slope $E_{\text{passivation}}$ similar to the modulus of elasticity E is shown in the case of the SGP model with $\mathcal{S}_{ijk}^{\text{dis}} \neq 0$, on the other hand, no elastic stress jump is observed in the case of the SGP model with $\mathcal{S}_{ijk}^{\text{dis}} = 0$. This result is exactly in agreement with the predictions in Fleck et al. (2014, 2015).

In the third part, an extensive parametric study is presented in terms of the effects of the dissipative length scale ℓ_{dis} , the onset point of passivation, the hardening material parameter h , the nonnegative rate sensitivity parameter m , the temperature sensitivity parameter n , the thermal material parameter T_y , the interfacial temperature sensitivity parameter n^I , and the interfacial thermal material parameter T_y^I on the stress-strain response for the two SGP cases, respectively. There are a number of conclusions worth mentioning here, namely: (1) the magnitude of the stress jump significantly increases as the dissipative length scale increases, on the other hand, the stress jump phenomenon disappears as the dissipative length scale comes closer to zero, (2) the slopes of the very first response E immediately after the passivation also increase as the dissipative length scales increase, (3) the stress jump phenomenon is highly correlated with the dissipative higher order microstress quantities $\mathcal{S}_{ijk}^{\text{dis}}$, in addition, the very first responses immediately after the passivation also make a substantial difference with varying passivation points, (4) the hardening material parameter h does not affect the stress jump significantly in the case of the SGP model with $\mathcal{S}_{ijk}^{\text{dis}} \neq 0$, (5) both the slope of the very first response immediately after the passivation and the corresponding magnitude of the stress jump substantially increase as the nonnegative rate sensitivity parameter m increases in the case of the SGP model with $\mathcal{S}_{ijk}^{\text{dis}} \neq 0$, (6) as the temperature-related parameters for the bulk such as the temperature sensitivity parameter n and the thermal material parameter T_y increase, the slope of the very first response

immediately after the passivation decreases and the corresponding magnitude of the stress jump increases in the case of the SGP model with $S_{ijk}^{\text{dis}} \neq 0$, (7) the slope of the very first response after the passivation, the corresponding magnitude of the stress jump increases and the variations along with the cases are shown more drastically by decreasing the temperature-related parameters for the interface, such as the interfacial temperature sensitivity parameter n^I and the interfacial thermal material parameter T_y^I in the case of the SGP model with $S_{ijk}^{\text{dis}} \neq 0$, and finally (8) no stress jump is observed in all cases with $S_{ijk}^{\text{dis}} = 0$.

Meanwhile, the plastic strain-dependent free energy accounting for the stored energy of cold work is derived in this chapter by assuming that the stored energy is related to the energy carried by dislocations. Accordingly, the variation of free energies and dissipation potentials during the plastic deformation are investigated with four different temperatures, i.e. 25 °C, 75 °C, 145 °C, and 218 °C. From the numerical results, it is shown that the stored energy of cold work has no temperature dependency; in addition, the stored energy of cold work tends to saturate after some critical point since the stored energy of cold work is proportional to the dislocation density that remains constant after the aforementioned critical point.

Lastly, the two-dimensional tension problem of the square plate (Voyiadjis and Song 2017) is solved to examine the mesh sensitivity of the model. The effects of the strain gradient and grain boundary are also studied. As expected, a strong mesh-dependence stress-strain behavior is observed in the case of no gradient effects, while the numerical results with the gradient effects show the mesh-independent behavior. The energetic hardening, the dissipative strengthening, the grain boundary strengthening, and the stress jump phenomena are well observed in common with the results from the one-dimensional simulation.

References

- Abaqus, *User's Manual (Version 6.12)* (Dassault Systemes Simulia Corp., Providence, 2012)
- E.C. Aifantis, *J. Eng. Mater-T. ASME* **106**, 4 (1984)
- K.E. Aifantis, J.R. Willis, *J. Mech. Phys. Solids* **53**, 5 (2005)
- B.D. Coleman, M.E. Gurtin, *J. Chem. Phys.* **47**, 2 (1967)
- N.A. Fleck, J.W. Hutchinson, *J. Mech. Phys. Solids* **49**, 10 (2001)
- N.A. Fleck, J.R. Willis, *J. Mech. Phys. Solids* **57**, 1 (2009a)
- N.A. Fleck, J.R. Willis, *J. Mech. Phys. Solids* **57**, 7 (2009b)
- N.A. Fleck, J.R. Willis, *Acta Mech. Sinica* **31**, 4 (2015)
- N.A. Fleck, J.W. Hutchinson, J.R. Willis, P. Roy, *Soc. A Math. Phys.* **470**, 2170 (2014)
- N.A. Fleck, J.W. Hutchinson, J.R. Willis, *J. Appl. Mech-T. ASME* **82**, 7 (2015)
- S. Forest, M. Amestoy, *C. R. Mecanique* **336**, 4 (2008)
- S. Forest, R. Sievert, *Acta Mech.* **160**, 1–2 (2003)
- P. Fredriksson, P. Gudmundson, *Mat. Sci. Eng. A Struct.* **400** (2005)
- P. Gudmundson, *J. Mech. Phys. Solids* **52**, 6 (2004)
- M.E. Gurtin, *J. Mech. Phys. Solids* **56**, 2 (2008)
- M.E. Gurtin, L. Anand, *J. Mech. Phys. Solids* **53**, 7 (2005)
- M.E. Gurtin, L. Anand, *J. Mech. Phys. Solids* **57**, 3 (2009)
- M.E. Gurtin, B.D. Reddy, *Contin. Mech. Thermodyn.* **21**, 3 (2009)

- S. Han, T. Kim, H. Lee, H. Lee, Electronics System-Integration Technology Conference, 2008. ESTC 2008, 2nd (2008)
- M.A. Haque, M.T.A. Saif, *Acta Mater.* **51**, 11 (2003)
- J.W. Hutchinson, *Acta Mech. Sinica* **28**, 4 (2012)
- M. Kuroda, V. Tvergaard, *Int. J. Plasticity* **26**, 4 (2010)
- D.B. Liu, Y.M. He, L. Shen, J. Lei, S. Guo, K. Peng, *Mat. Sci. Eng. A-Struct.* **647** (2015)
- M.A. Meyers, K.K. Chawla, *Mechanical Behavior of Materials* (Cambridge University Press, 2009), Cambridge, United Kingdom
- F. Mollica, K.R. Rajagopal, A.R. Srinivasa, *Int. J. Plast.* **17**, 8 (2001)
- H.B. Muhlhaus, E.C. Aifantis, *Int. J. Solids Struct.* **28**, 7 (1991)
- L. Nicola, Y. Xiang, J.J. Vlassak, E. Van der Giessen, A. Needleman, *J. Mech. Phys. Solids* **54**, 10 (2006)
- P. Rosakis, A.J. Rosakis, G. Ravichandran, J. Hodowany, *J. Mech. Phys. Solids* **48**, 3 (2000)
- S.S. Shishvan, L. Nicola, E. Van der Giessen, *J. Appl. Phys.* **107**, 9 (2010)
- S.S. Shishvan, S. Mohammadi, M. Rahimian, E. Van der Giessen, *Int. J. Solids Struct.* **48**, 2 (2011)
- G.Z. Voyiadjis, B. Deliktas, *Int. J. Eng. Sci.* **47**, 11–12 (2009)
- G.Z. Voyiadjis, D. Faghihi, *Int. J. Plast.* 30–31 (2012)
- G.Z. Voyiadjis, D. Faghihi, *J. Eng. Mater-T. ASME* **136**, 4 (2014)
- G.Z. Voyiadjis, Y. Song, *Philos. Mag.* **97**, 5 (2017)
- G.Z. Voyiadjis, D. Faghihi, Y.D. Zhang, *Int. J. Solids Struct.* **51**, 10 (2014)
- G.Z. Voyiadjis, Y. Song, T. Park, *J. Eng. Mater-T. ASME* **139**, 2 (2017)
- Y. Xiang, J.J. Vlassak, *Acta Mater.* **54**, 20 (2006)
- Y. Xiang, X. Chen, J.J. Vlassak, *J. Mater. Res.* **20**, 9 (2005)



Micropolar Crystal Plasticity

17

J. R. Mayeur, D. L. McDowell, and Samuel Forest

Contents

Introduction	596
Classical Single Crystal Plasticity	597
Finite Deformation Kinematics	597
Thermodynamics	597
Deformation Incompatibility and the GND Density Tensor	600
Micropolar Single Crystal Plasticity	602
Finite Deformation Theory	602
Linear Deformation Theory	605
Relationship to Slip Gradient Theory	611
Applications: Comparison to 2D Discrete Dislocation Dynamics Simulations	618
Constrained Shear of Thin Films	620
Pure Bending of Thin Films: Single Slip	623
Pure Bending of Thin Films: Double Slip	630
Simple Shear of a Metal Matrix Composite	633
Conclusions	639
References	640

J. R. Mayeur

Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM, USA

e-mail: jason.mayeur@gmail.com

D. L. McDowell

Woodruff School of Mechanical Engineering, School of Materials Science and Engineering,
Georgia Institute of Technology, Atlanta, GA, USA

e-mail: david.mcdowell@me.gatech.edu

S. Forest (✉)

Centre des Materiaux, Mines ParisTech CNRS, PSL Research University, CNRS UMR 7633,
Evry Cedex, Paris, France

e-mail: samuel.forest@mines-paristech.fr

© Springer Nature Switzerland AG 2019

G. Z. Voyiadjis (ed.), *Handbook of Nonlocal Continuum Mechanics for Materials and Structures*, https://doi.org/10.1007/978-3-319-58729-5_48

595

Abstract

This chapter considers advances over the past 15 years achieved by the authors and coworkers on generalized crystal plasticity to address size and configuration effects in dislocation plasticity at the micron scale. The specific approaches addressed here focus on micropolar and micromorphic theories rather than adopting strain gradient theory as the starting point, as motivated by the pioneering ideas of Eringen (Eringen and Suhubi 1964; Eringen and Claus Jr 1969; Eringen 1999). It is demonstrated with examples that for isotropic elasticity and specific sets of slip systems, a dislocation-based formulation of micropolar or micromorphic type provides results comparable to discrete dislocation dynamics and has much in common with the structure of Gurtin's slip gradient theory (Gurtin 2002; Gurtin et al. 2007).

Keywords

Micropolar · Strain gradient · GNDs · Crystal plasticity · Finite elements

Introduction

The collective behavior of dislocations in a single crystal can be described by means of the continuum theory of dislocations. The material volume element is assumed to contain a suitable density of dislocations for the continuum theory of dislocations to be applicable. Nonhomogeneous plastic deformations induce material and lattice incompatibilities that are resolved by a suitable distribution of the dislocation density tensor field which can be interpreted as a second rank statistical mean for a population of arbitrary dislocations inside a material volume element (Kröner 1969; Cermelli and Gurtin 2001). Nye's fundamental relation linearly connects the dislocation density tensor to the lattice curvature field of the crystal. This fact has prompted many authors to treat a continuously dislocated crystal as a Cosserat continuum (Günther 1958; Kröner 1963; Schäfer 1969; Forest et al. 2000). The Cosserat approach records only the lattice curvature of the crystal but neglects the effect of the rotational part of the elastic strain tensor, which is a part of the total dislocation density tensor (Cordero et al. 2010). Full account of plastic incompatibility is taken in strain gradient plasticity theories, starting from the original work by Aifantis (1984) up to the work of Gurtin (2002). Formulation of crystal plasticity within the micromorphic framework is more recent and was suggested by Clayton et al. (2005) for a large spectrum of crystal defects, including point defects and disclinations. Limiting the discussion to dislocation density tensor effects, also called geometrically necessary dislocation (GND) effects, Cordero et al. (2010) showed, within a small deformation setting, how the micromorphic model can be used to predict grain and precipitate size effects in laminate crystalline materials. In particular, the micromorphic model is shown to deliver more general scaling laws than conventional strain gradient plasticity. These models represent extensions of the conventional crystal plasticity theory (cf. Teodosiu and Sidoroff

1976) that accounts for single crystal hardening and lattice rotation but does not incorporate the effect of the dislocation density tensor.

The layout of the chapter is as follows. Classical single crystal plasticity theory is first recalled in section “[Finite Deformation Kinematics](#)” including the thermodynamical framework and the definition of the continuum dislocation density tensor. The Cosserat generalization of crystal plasticity is presented in section “[Micropolar Single Crystal Plasticity](#)” with its relation to strain gradient plasticity. The section “[Applications: Comparison to 2D Discrete Dislocation Dynamics Simulations](#)” provides applications including a comparison between Cosserat constitutive laws and results from discrete dislocation dynamics in the case of constrained thin films in shear and bending.

Classical Single Crystal Plasticity

Finite Deformation Kinematics

The classical theory of finite deformation single crystal plasticity is based on a multiplicative decomposition of the deformation gradient into elastic and plastic parts,

$$\mathbf{F} = \mathbf{F}^e \mathbf{F}^p \quad (1)$$

where \mathbf{F}^p describes the plastic deformation of the continuum that leaves the underlying lattice vectors unaltered and \mathbf{F}^e describes the elastic stretching and rotation of the lattice relative to this intermediate, isoclinic configuration. The deformation gradient maps infinitesimal vectors from the reference to current configuration and can be expressed in terms of the referential gradient of the displacement field as

$$d\mathbf{x} = \mathbf{F}d\mathbf{X}, \quad \mathbf{F} = \mathbf{1} + \mathbf{H}, \quad \mathbf{H} = \mathbf{u}\nabla_0 \quad (2)$$

\mathbf{H} is the distortion (or displacement gradient) tensor and has been introduced for use in subsequent sections. The theory is completed by supplying constitutive prescriptions for the relationships between \mathbf{F}^e and the Cauchy stress and to provide an evolution equation for \mathbf{F}^p (and any associated internal state variables) consistent with thermodynamics.

Thermodynamics

The standard nonpolar mechanical balance laws (neglecting inertial terms) are given in the current configuration as

$$\boldsymbol{\sigma} \cdot \nabla + \mathbf{f} = \mathbf{0}, \quad \boldsymbol{\sigma} = \boldsymbol{\sigma}^T \quad (3)$$

with the associated traction boundary condition, $\mathbf{t} = \boldsymbol{\sigma} \mathbf{n}$, where $\boldsymbol{\sigma}$ is the Cauchy stress tensor, \mathbf{f} is a body force vector, \mathbf{t} is the traction vector, and \mathbf{n} is the unit normal to the external part of the boundary where tractions are specified. The total energy balance can be expressed as

$$\rho \dot{U} = \boldsymbol{\sigma} : \mathbf{D} - \nabla \cdot \mathbf{q} \quad (4)$$

where ρ is the current density, U is the specific internal energy, \mathbf{D} is the rate of deformation tensor, and \mathbf{q} is the heat flux vector. The entropy inequality which will be used to derive the state equations and guide the construction of plastic evolution equations is given as

$$\rho \dot{\eta} + \nabla \cdot \left(\frac{\mathbf{q}}{T} \right) \geq 0 \quad (5)$$

where η is the specific entropy, and T is temperature. Using the state relation $\psi = U - T\eta$, Eqs. (4) and (5) can be combined to obtain the Clausius-Duhem inequality

$$\boldsymbol{\sigma} : \mathbf{D} - \rho (\dot{\psi} + \dot{T}\eta) - \frac{1}{T} \mathbf{q} \cdot \nabla T \geq 0 \quad (6)$$

where the intrinsic (Δ^{intr}) and thermal (Δ^{th}) dissipation are defined as

$$\Delta^{intr} = \boldsymbol{\sigma} : \mathbf{D} - \rho (\dot{\psi} + \dot{T}\eta), \quad \Delta^{th} = \frac{1}{T} \mathbf{q} \cdot \nabla T \quad (7)$$

Next, we derive an expression for the stress power in the intermediate configuration by using the relationship between the Cauchy stress and the second Piola-Kirchhoff stress with respect to the intermediate configuration $\tilde{\mathbf{S}} = J^e \mathbf{F}^{e-1} \boldsymbol{\sigma} \mathbf{F}^{e-T}$ and by writing the velocity gradient in terms of the multiplicative decomposition, i.e.,

$$\mathbf{L} = \dot{\mathbf{F}} \mathbf{F}^{-1} = \dot{\mathbf{F}}^e \mathbf{F}^{e-1} + \mathbf{F}^e \dot{\mathbf{F}}^p \mathbf{F}^{p-1} \mathbf{F}^{e-1} \quad (8)$$

such that

$$\frac{1}{\rho} \boldsymbol{\sigma} : \mathbf{L} = \frac{1}{\tilde{\rho}} \left(\tilde{\mathbf{S}} : \dot{\tilde{\mathbf{E}}}^e + \boldsymbol{\Pi}^M : \tilde{\mathbf{L}}^p \right) \quad (9)$$

where $\tilde{\rho}$ is the density in the intermediate configuration, $\tilde{\mathbf{E}}^e$ is the elastic Green-Lagrange strain, $\boldsymbol{\Pi}^M$ is the Mandel stress tensor, and $\tilde{\mathbf{L}}^p$ is the plastic velocity gradient in the intermediate configuration which have the following definitions

$$\tilde{\mathbf{E}}^e = \frac{1}{2} (\mathbf{F}^{eT} \mathbf{F}^e - \mathbf{1}), \quad \boldsymbol{\Pi}^M = \mathbf{F}^{eT} \mathbf{F}^e \tilde{\mathbf{S}}, \quad \tilde{\mathbf{L}}^p = \dot{\mathbf{F}}^p \mathbf{F}^{p-1}. \quad (10)$$

Equation (9) identifies the appropriate power-conjugate variables in the intermediate configuration, from which it can be seen that $\tilde{\mathbf{E}}^e$ is the strain measure power-conjugate to $\tilde{\mathbf{\Pi}}^e$ and that $\mathbf{\Pi}^M$ is power-conjugate to $\tilde{\mathbf{L}}^p$. If we now introduce the Helmholtz free energy, $\psi = \hat{\psi}(\tilde{\mathbf{E}}^e, T, \zeta^\alpha)$, the following expression is obtained by inserting the right-hand side of Eq. (9) along with the chain rule expression of the time derivative of ψ into the Clausius-Duhem inequality

$$\left(\mathbf{\Pi}^e - \tilde{\rho} \frac{\partial \psi}{\partial \tilde{\mathbf{E}}^e} \right) : \dot{\tilde{\mathbf{E}}}^e - \tilde{\rho} \left(\eta + \frac{\partial \psi}{\partial T} \right) \dot{T} + \mathbf{\Pi}^M : \tilde{\mathbf{L}}^p - \tilde{\rho} \sum_{\alpha} \frac{\partial \psi}{\partial \zeta^{\alpha}} * \dot{\zeta}^{\alpha} \geq 0. \quad (11)$$

where ζ^α is a set of internal state variables. Here, “*” is an appropriate scalar product operator for the rank of tensor ζ^α . The state laws are then deduced as

$$\mathbf{\Pi}^e = \tilde{\rho} \frac{\partial \psi}{\partial \tilde{\mathbf{E}}^e}, \quad \eta = - \frac{\partial \psi}{\partial T}. \quad (12)$$

and the residual intrinsic dissipation is then expressed as

$$\Delta^{intr} = \mathbf{\Pi}^M : \tilde{\mathbf{L}}^p - \sum_{\alpha} r^{\alpha} * \dot{\zeta}^{\alpha} \geq 0, \quad r^{\alpha} = \tilde{\rho} \frac{\partial \psi}{\partial \zeta^{\alpha}}. \quad (13)$$

Therefore, thermodynamically consistent evolution equations for $\tilde{\mathbf{L}}^p$ and $\dot{\zeta}^{\alpha}$ may be derived by introducing a convex dissipation potential $\Omega = \hat{\Omega}(\mathbf{\Pi}^M, r^{\alpha})$ such that

$$\tilde{\mathbf{L}}^p = \frac{\partial \Omega}{\partial \mathbf{\Pi}^M}, \quad \dot{\zeta}^{\alpha} = - \frac{\partial \Omega}{\partial r^{\alpha}} \quad (14)$$

Representative functional forms for $\hat{\psi}$ and $\hat{\Omega}$ are given for completeness below. It is typical to use a free energy that is quadratic with respect to its arguments, i.e.,

$$\tilde{\rho} \psi(\tilde{\mathbf{E}}^e, \zeta^{\alpha}) = \frac{1}{2} \tilde{\mathbf{E}}^e : \tilde{\mathbf{C}} : \tilde{\mathbf{E}}^e + \frac{1}{2} \sum_{\alpha, \beta} a^{\alpha\beta} \zeta^{\alpha} \zeta^{\beta} \quad (15)$$

where $\tilde{\mathbf{C}}$ is the fourth-order elasticity tensor in the intermediate configuration and $a^{\alpha\beta}$ is positive definite interaction matrix that describes the coupling between the ζ^{α} . Likewise, a typical power law potential for the inelastic evolution equations is given as

$$\Omega(\mathbf{\Pi}^M, r^{\alpha}) = \frac{\dot{\gamma}_0}{n+1} \sum_{\alpha} g^{\alpha} \left\langle \frac{\mathcal{F}^{\alpha}}{g^{\alpha}} \right\rangle^{n+1} \quad (16)$$

where n is the power law exponent, g^{α} is viscous stress, $\dot{\gamma}_0$ a typical strain rate parameter, and \mathcal{F}^{α} is a yield function. The brackets $\langle \bullet \rangle = \text{Max}(\bullet, 0)$ have been

introduced. The yield function is defined in terms of the resolved shear stress, τ^α , and the energetic flow resistance, r^α , as

$$\mathcal{F}^\alpha = \tau^\alpha - (r_0 + r^\alpha), \quad \tau^\alpha = \boldsymbol{\Pi}^M : (\tilde{\mathbf{s}}^\alpha \otimes \tilde{\mathbf{n}}^\alpha) \quad (17)$$

The slip direction vector and the normal to the slip plane for slip system α are respectively denoted by $\tilde{\mathbf{s}}^\alpha$ and $\tilde{\mathbf{n}}^\alpha$ in the undistorted lattice configuration.

Deformation Incompatibility and the GND Density Tensor

Finite Deformation Kinematics

For a classical Cauchy continuum, compatibility of the displacement field requires that the curl of the deformation gradient vanishes, i.e.,

$$\oint_c d\mathbf{x} = \oint_c \mathbf{F} d\mathbf{X} = \mathbf{0} \Rightarrow \mathbf{F} \times \nabla_0 = \mathbf{0} \quad (18)$$

During an inhomogeneous elastic-plastic deformation, the elastic and plastic deformation maps are not compatible and can be used to quantify the heterogeneity of the deformation field in terms of the net Burgers on the intermediate configuration, i.e.,

$$\tilde{\mathbf{b}} = \oint_{\tilde{c}} d\tilde{\mathbf{x}} = \oint_c \mathbf{F}^{e-1} d\mathbf{x} = \oint_c \mathbf{F}^p d\mathbf{X} \quad (19)$$

Making use of Stokes' theorem, the last two expressions in Eq. (19) can be expressed in terms of surface integrals as

$$\tilde{\mathbf{b}} = \int_s \mathbf{A}^e \mathbf{n} da = \oint_S \mathbf{A}^p \mathbf{N} dA \quad (20)$$

where $\mathbf{A}^e = \mathbf{F}^{e-1} \times \nabla$ and $\mathbf{A}^p = \mathbf{F}^p \times \nabla_0$ are the corresponding two-point geometrically necessary dislocation density tensors that map from current to intermediate and reference to intermediate configurations, respectively.

There have been many works in the last two decades focused on incorporating the effects of GNDs into crystal plasticity modeling frameworks. It is beyond the scope of this chapter to attempt to review the myriad ways in which these extensions are carried out. The vast majority of generalized crystal plasticity models that incorporate the effects of GNDs do so by computing them from the plastic slip gradients (so-called slip gradient theories) rather than via gradients of \mathbf{F}^{e-1} . While the connection between slip gradients and GND densities can be established in a finite deformation context (see Kuroda and Tvergaard 2008), in this section, the presentation is limited to the linearized kinematic setting for ease of presentation

and direct connection with the pioneering work of Nye (1953). The subsequent developments closely follow the presentation of Arsenlis and Parks (1999).

Linearized Kinematics

In the case of linearized kinematics, the multiplicative decomposition of the deformation gradient is replaced with an additive decomposition of the distortion tensor, i.e., $\mathbf{H} = \mathbf{H}^e + \mathbf{H}^p$ and Eq. (18) may conveniently be rewritten as

$$\mathbf{H} \times \nabla = 0 \quad (21)$$

The continuum GND density tensor, $\mathbf{A} \approx \mathbf{A}^e \approx \mathbf{A}^p$, can be equivalently expressed in terms of either the elastic or plastic distortion as

$$\mathbf{A} = -\mathbf{H}^e \times \nabla = \mathbf{H}^p \times \nabla \quad (22)$$

The lattice torsion-curvature is defined as the gradient of the lattice rotation vector (Nye 1953), i.e.,

$$\boldsymbol{\kappa} = \phi \nabla \quad (23)$$

where ϕ is given by

$$\phi = -\frac{1}{2} \boldsymbol{\epsilon} : \text{skw}(\mathbf{H}^e) \quad (24)$$

where “skw” is the skew operator providing the antisymmetric part of the tensor. The notation $\boldsymbol{\epsilon}$ is used for the permutation tensor. Combining Eqs. (22), (23), and (24) the lattice torsion-curvature may be expressed as

$$\boldsymbol{\chi} = \underbrace{-\mathbf{A}^T + \frac{1}{2} \text{tr}(\mathbf{A}) \mathbf{1}}_{\text{Nye curvature, } \boldsymbol{\kappa}} - (\boldsymbol{\varepsilon}^e \times \nabla)^T \quad (25)$$

where $\boldsymbol{\varepsilon}^e$ is the elastic strain tensor which is symmetric. As indicated in Eq. (25), the first two terms represent Nye’s original torsion-curvature tensor since it was assumed that $\mathbf{H}^e \approx \text{skw}(\mathbf{H}^e)$.

Next, we seek an expression for \mathbf{A} in terms of slip gradients which can be obtained from taking the curl of the plastic distortion. First, note that the discrete version of Nye’s tensor, \mathbf{A}^d , for a population of straight edge and screw dislocations can be written as

$$\mathbf{A}^d = b \sum_{\alpha} \underbrace{(\varrho_{\perp}^{\alpha} - \varrho_{\top}^{\alpha})}_{\varrho_{G\perp}^{\alpha}} \mathbf{s}^{\alpha} \otimes \mathbf{t}^{\alpha} + \underbrace{(\varrho_{\odot}^{\alpha} - \varrho_{\otimes}^{\alpha})}_{\varrho_{G\odot}^{\alpha}} \mathbf{s}^{\alpha} \otimes \mathbf{s}^{\alpha} \quad (26)$$

where ϱ_{\perp}^{α} is the positive edge dislocation density, ϱ_{\top}^{α} is the negative edge dislocation density, ϱ_{\odot}^{α} is the positive screw dislocation density, $\varrho_{\otimes}^{\alpha}$ is the negative screw dislocation density, $\varrho_{G\perp}^{\alpha}$ is the edge GND density, and $\varrho_{G\odot}^{\alpha}$ is the screw GND density. The continuous expression of Nye's GND density tensor is then obtained as

$$\begin{aligned} \mathbf{A} &= \mathbf{H}^p \times \nabla \\ &= \sum_{\alpha} \gamma^{\alpha} \mathbf{s}^{\alpha} \otimes \mathbf{n}^{\alpha} \times \nabla \\ &= \sum_{\alpha} (-\nabla \gamma^{\alpha} \cdot \mathbf{s}^{\alpha}) \mathbf{s}^{\alpha} \otimes \mathbf{t}^{\alpha} + (\nabla \gamma^{\alpha} \cdot \mathbf{t}^{\alpha}) \mathbf{s}^{\alpha} \otimes \mathbf{s}^{\alpha} \end{aligned} \quad (27)$$

Comparing these two expressions for Nye's tensor, it may be shown that the continuum GND densities are given by slip gradients projected in the glide directions for the respective dislocation populations, i.e.,

$$\varrho_{G\perp}^{\alpha} = -b^{-1} \nabla \gamma^{\alpha} \cdot \mathbf{s}^{\alpha}, \quad \varrho_{G\odot}^{\alpha} = b^{-1} \nabla \gamma^{\alpha} \cdot \mathbf{t}^{\alpha} \quad (28)$$

An expression relating Nye's torsion-curvature tensor, $\boldsymbol{\kappa}$, to the GND densities is obtained by inserting Eq. (27) into the first two terms in Eq. (25)

$$\boldsymbol{\kappa} = -b \sum_{\alpha} \left[\varrho_{G\perp}^{\alpha} \mathbf{t}^{\alpha} \otimes \mathbf{s}^{\alpha} + \varrho_{G\odot}^{\alpha} \left(\mathbf{s}^{\alpha} \otimes \mathbf{s}^{\alpha} - \frac{1}{2} \mathbf{1} \right) \right] \quad (29)$$

We will revisit the expression given in Eq. (29) when motivating constitutive equations in subsequent sections.

Micropolar Single Crystal Plasticity

Finite Deformation Theory

The presentation is given here within the finite deformation framework before returning to the linearized case.

Kinematics

A micropolar continuum is a generalized continua with extra rotational degrees of freedom. Considering two sets of vector triads attached to each material point in the reference configuration, \mathbf{X} , there is an independent mapping of the two sets of vectors to the current configuration such that

$$d\mathbf{x}_i = \mathbf{F}d\mathbf{X}_i, \quad \mathbf{d}_i = \overline{\mathbf{R}}\mathbf{D}_i \quad \forall i = 1, 3 \quad (30)$$

where \mathbf{F} is the usual deformation gradient (see Eq. (2)) and $\overline{\mathbf{R}}$ is the two-point tensor that maps the microstructure triad in the reference configuration, \mathbf{D}_i to its image in

the current configuration, \mathbf{d}_i . $\bar{\mathbf{R}}$ is the micropolar rotation tensor; the overbar is used to distinguish it from the rotational part of the deformation gradient obtained via the polar decomposition, i.e., $\mathbf{F} = \mathbf{R}\mathbf{U}$. The micropolar rotation tensor is a proper orthogonal tensor such that the following relations hold:

$$\bar{\mathbf{R}} \bar{\mathbf{R}}^T = \mathbf{1}, \quad \bar{\mathbf{R}}(\mathbf{X}, t = 0) = \mathbf{1}, \quad \det(\bar{\mathbf{R}}) = 1 \quad (31)$$

The micropolar rotation field can also be expressed in terms of the axial vector field, $\bar{\boldsymbol{\phi}}$, via

$$\bar{\mathbf{R}} = \exp(\bar{\boldsymbol{\Phi}}), \quad \bar{\Phi}_{ij} = -\varepsilon_{ijk} \bar{\phi}_k \quad (32)$$

where $\bar{\boldsymbol{\Phi}}$ is the skew symmetric tensor associated with the axial vector $\bar{\boldsymbol{\phi}}$. The three components of $\bar{\boldsymbol{\phi}}$ along with the displacement field, \mathbf{u} , comprise the six independent degrees of freedom for the micropolar continuum. It has been shown by Eringen and Suhubi (1964) that a suitable set of invariant Lagrangian strain measures for the micropolar continuum may be defined as

$$\bar{\mathbf{U}} = \bar{\mathbf{R}}^T \mathbf{F}, \quad \boldsymbol{\Gamma} = \bar{\mathbf{R}}^T (\bar{\mathbf{R}} \nabla_0) \quad (33)$$

where $\bar{\mathbf{U}}$ is called the relative deformation tensor and $\boldsymbol{\Gamma}$ is the third-rank wryness (or torsion-curvature tensor). The Lagrangian micropolar strain is defined in terms of the relative distortion tensor as $\bar{\mathbf{E}} = \bar{\mathbf{U}} - \mathbf{1}$. Due to the antisymmetry with respect to its first two indices, it is convenient to express the torsion-curvature as a second-order tensor, i.e.,

$$\bar{\mathbf{K}} = -\frac{1}{2} \boldsymbol{\epsilon} : \boldsymbol{\Gamma} \quad (34)$$

The rates of the micropolar strain and torsion-curvature tensors are related to the velocity, $\mathbf{v} = \dot{\mathbf{u}}$, and the microstructural angular velocity, $\bar{\boldsymbol{\Omega}} = \dot{\bar{\mathbf{R}}} \bar{\mathbf{R}}^T$, as

$$\dot{\bar{\mathbf{U}}} \bar{\mathbf{U}}^{-1} = \bar{\mathbf{R}}^T (\mathbf{L} - \bar{\boldsymbol{\Omega}}) \bar{\mathbf{R}} \quad (35)$$

$$\dot{\bar{\mathbf{K}}} \bar{\mathbf{U}}^{-1} = \bar{\mathbf{R}}^T (\bar{\boldsymbol{\omega}} \nabla) \bar{\mathbf{R}} \quad (36)$$

where $\mathbf{L} = \dot{\mathbf{F}} \mathbf{F}^{-1}$ is the velocity gradient and $\bar{\boldsymbol{\omega}} = -1/2 \boldsymbol{\epsilon} : \bar{\boldsymbol{\Omega}}$ is the axial vector associated with $\bar{\boldsymbol{\Omega}}$. From Eqs. (35) and (36), it is clear that the rate expressions on the LHS are pull-backs from the current configuration via the microrotation.

Thermodynamics

The mechanical balance laws under static equilibrium for the micropolar continuum are

$$\nabla \cdot \boldsymbol{\sigma}^T + \mathbf{f} = 0, \quad \nabla \cdot \mathbf{m}^T - \boldsymbol{\epsilon} : \boldsymbol{\sigma} + \mathbf{c} = 0 \quad (37)$$

where $\boldsymbol{\sigma}$ is the unsymmetric Cauchy force stress tensor, \mathbf{f} is a body force vector, \mathbf{m} is the couple stress tensor, and \mathbf{c} is a body couple vector. These balance laws can be deduced using the principle of virtual work (Germain 1973; Forest and Sievert 2003). The energy balance and Clausius-Duhem inequality for the micropolar continuum are given, respectively, as

$$\rho \dot{U} = \boldsymbol{\sigma} : (\mathbf{L} - \overline{\boldsymbol{\Omega}}) + \mathbf{m} : \overline{\boldsymbol{\omega}} \nabla - \nabla \cdot \mathbf{q} \quad (38)$$

$$\hat{\boldsymbol{\sigma}} : \left(\dot{\overline{\mathbf{U}}} \overline{\mathbf{U}}^{-1} \right) + \hat{\mathbf{m}} : \left(\dot{\overline{\mathbf{K}}} \overline{\mathbf{U}}^{-1} \right) - \rho (\dot{\psi} + \eta \dot{T}) - \frac{1}{T} \mathbf{q} \cdot \nabla T \geq 0 \quad (39)$$

where $\hat{\boldsymbol{\sigma}}$ and $\hat{\mathbf{m}}$ are the Cauchy and couple stress tensors pulled back to the reference configuration via $\overline{\mathbf{R}}$, i.e.,

$$\hat{\boldsymbol{\sigma}} = \overline{\mathbf{R}}^T \boldsymbol{\sigma} \overline{\mathbf{R}}, \quad \hat{\mathbf{m}} = \overline{\mathbf{R}}^T \mathbf{m} \overline{\mathbf{R}} \quad (40)$$

The specific form of Eq. (39) was obtained by using the kinematic relations given in Eqs. (35) and (36). Let us first consider the case of finite micropolar thermoelasticity and assume that $\psi = \hat{\psi}(\overline{\mathbf{U}}, \overline{\mathbf{K}}, T)$. Taking the time derivative of ψ and inserting into Eq. (39) leads to

$$\left(\hat{\boldsymbol{\sigma}} \overline{\mathbf{U}}^{-T} - \rho \frac{\partial \psi}{\partial \overline{\mathbf{U}}} \right) : \dot{\overline{\mathbf{U}}} + \left(\hat{\mathbf{m}} \overline{\mathbf{U}}^{-T} - \rho \frac{\partial \psi}{\partial \overline{\mathbf{K}}} \right) : \dot{\overline{\mathbf{K}}} - \rho \left(\eta + \frac{\partial \psi}{\partial T} \right) \dot{T} \geq 0 \quad (41)$$

Therefore, the state laws for the micropolar material are

$$\hat{\boldsymbol{\sigma}} = \rho \frac{\partial \psi}{\partial \overline{\mathbf{U}}} \overline{\mathbf{U}}^T, \quad \hat{\mathbf{m}} = \rho \frac{\partial \psi}{\partial \overline{\mathbf{K}}} \overline{\mathbf{U}}^T, \quad \eta = -\frac{\partial \psi}{\partial T} \quad (42)$$

Next we must introduce elastic-plastic decompositions for $\overline{\mathbf{U}}$ and $\overline{\mathbf{K}}$. We start with the natural assumption that the hyperelastic relations will have the same form as Eq. (42) with respect to the elastic deformation measures, i.e.,

$$\hat{\boldsymbol{\sigma}} = \rho \frac{\partial \psi}{\partial \overline{\mathbf{U}}^e} \overline{\mathbf{U}}^{eT}, \quad \hat{\mathbf{m}} = \rho \frac{\partial \psi}{\partial \overline{\mathbf{K}}^e} \overline{\mathbf{U}}^{eT} \quad (43)$$

Consistent with this assumption, it can be shown (Sievert et al. 1998) that the appropriate decompositions are given as

$$\bar{\mathbf{U}} = \bar{\mathbf{U}}^e \bar{\mathbf{U}}^p, \quad \bar{\mathbf{K}} = \bar{\mathbf{K}}^e \bar{\mathbf{U}}^p + \bar{\mathbf{K}}^p \quad (44)$$

Substituting the rate forms of these decompositions back into the Clausius-Duhem inequality with $\psi = \hat{\psi}(\bar{\mathbf{U}}^e, \bar{\mathbf{K}}^e, T, \zeta^\alpha)$ yields

$$\Sigma : \left(\dot{\bar{\mathbf{U}}}^p \bar{\mathbf{U}}^{p-1} \right) + \mathbf{M} : \left(\dot{\bar{\mathbf{K}}}^p \bar{\mathbf{U}}^{p-1} \right) + \rho \sum_{\alpha} \frac{\partial \psi}{\partial \zeta^\alpha} * \dot{\zeta}^\alpha \geq 0 \quad (45)$$

where the driving forces for the plastic evolution equations are identified as

$$\Sigma = \bar{\mathbf{U}}^{eT} \hat{\boldsymbol{\sigma}} \bar{\mathbf{U}}^{e-T} + \bar{\mathbf{K}}^{eT} \hat{\mathbf{m}} \bar{\mathbf{U}}^{e-T}, \quad \mathbf{M} = \hat{\mathbf{m}} \bar{\mathbf{U}}^{e-T} \quad (46)$$

Specific constitutive equations must be provided for the free energy and plastic evolution equations to complete the formulation. The procedure for doing so can be undertaken in analogous fashion to what is done in the classical theory. However, there is some additional flexibility that is afforded in constructing the plastic evolution equations in this case compared to the classical theory (see Forest and Sievert (2003) for an in-depth discussion of single vs. multi-criterion flow rules). Rather than narrowing to specific constitutive choices in the current finite deformation context, we will discuss these issues within the small deformation framework in the subsequent sections. The principles guiding constitutive equation development are the same for both finite and infinitesimal deformations and we choose to discuss these aspects with respect to the theories used in the numerical simulations appearing later in the chapter. A more complete exposition of constitutive equation development in the finite deformation context has been given elsewhere (Forest et al. 1997; Forest 2012).

Linear Deformation Theory

Linearized Kinematics

In the case of small deformations and rotations, the following notation is introduced for the micropolar strain, $\bar{\boldsymbol{\varepsilon}}$, and torsion-curvature, $\bar{\boldsymbol{\kappa}}$:

$$\bar{\mathbf{E}} = \bar{\mathbf{U}} - \mathbf{1} \approx \mathbf{u} \nabla - \bar{\boldsymbol{\Phi}} =: \bar{\boldsymbol{\varepsilon}} = \bar{\boldsymbol{\varepsilon}}^e + \bar{\boldsymbol{\varepsilon}}^p \quad (47)$$

$$\bar{\mathbf{K}} = -\frac{1}{2} \boldsymbol{\epsilon} : \boldsymbol{\Gamma} \approx \bar{\boldsymbol{\phi}} \nabla =: \bar{\boldsymbol{\kappa}} = \bar{\boldsymbol{\kappa}}^e + \bar{\boldsymbol{\kappa}}^p \quad (48)$$

As indicated in Eqs. (47) and (48), we assume an additive elastic-plastic decomposition of the strain and torsion-curvature tensors. Note that the symmetric part of $\bar{\boldsymbol{\varepsilon}}$ is the classical small strain tensor: $\text{sym}(\bar{\boldsymbol{\varepsilon}}) = \boldsymbol{\epsilon} = \text{sym}(\mathbf{H})$, and the skew-symmetric part is a measure of the difference between the continuum rotation,

$\mathbf{w} = \text{skw}(\mathbf{H})$, and microrotation: $\text{skw}(\bar{\boldsymbol{\varepsilon}}) = \mathbf{w} - \bar{\boldsymbol{\Phi}}$. The additive decomposition of the distortion tensor, \mathbf{H} , into elastic and plastic parts is assumed (see section “[Linearized Kinematics](#)”) and the micropolar plastic strain is defined to be equal to the plastic distortion such that the evolution equation of $\bar{\boldsymbol{\varepsilon}}^p$ has the form

$$\dot{\bar{\boldsymbol{\varepsilon}}}^p = \sum_{\alpha} \dot{\gamma}^{\alpha} \mathbf{s}^{\alpha} \otimes \mathbf{n}^{\alpha} \quad (49)$$

Therefore, elastic micropolar strain is defined as

$$\bar{\boldsymbol{\varepsilon}}^e = \mathbf{H}^e - \bar{\boldsymbol{\Phi}} = \underbrace{\boldsymbol{\varepsilon}^e}_{\text{sym}(\bar{\boldsymbol{\varepsilon}}^e)} + \underbrace{(\mathbf{w}^e - \bar{\boldsymbol{\Phi}})}_{\text{skw}(\bar{\boldsymbol{\varepsilon}}^e)} \quad (50)$$

Equation (50) shows that the skew-symmetric part of the micropolar elastic strain is just the difference between the lattice rotation embodied by \mathbf{w}^e and the microrotation. Using the relationship between the lattice torsion-curvature and GNDs presented in section “[Linearized Kinematics](#)” as motivation, Forest et al. (1997) proposed a micropolar plastic torsion-curvature evolution equation of the form (note there is a sign convention difference between the screw GND term presented here and what was originally proposed in that work):

$$\dot{\bar{\boldsymbol{\kappa}}}^p = \sum_{\alpha} \left[\frac{\dot{\varphi}_{\perp}^{\alpha}}{L_{\perp}} \mathbf{t}^{\alpha} \otimes \mathbf{s}^{\alpha} + \frac{\dot{\varphi}_{\odot}^{\alpha}}{L_{\odot}} \left(\mathbf{s}^{\alpha} \otimes \mathbf{s}^{\alpha} - \frac{1}{2} \mathbf{1} \right) \right]. \quad (51)$$

Here, $\dot{\varphi}_{\perp}^{\alpha}$ and L_{\perp}^{α} are the plastic rotation rate and plastic length scale associated with edge GNDs and $\dot{\varphi}_{\odot}^{\alpha}$ and L_{\odot}^{α} are the analogous quantities for screw GNDs. Comparing Eqs. (29) and (51) reveals the relationship between the GND densities and the micropolar plastic torsion-curvature parameters, namely

$$\dot{\varrho}_{G_{\perp}}^{\alpha} = -\frac{\dot{\varphi}_{\perp}^{\alpha}}{bL_{\perp}}, \quad \dot{\varrho}_{G_{\odot}}^{\alpha} = -\frac{\dot{\varphi}_{\odot}^{\alpha}}{bL_{\odot}} \quad (52)$$

The introduced length scales are expected to be in the range of micron and submicron sizes as illustrated in the examples provided in this chapter.

Thermodynamics at Small Strains

Let us now revisit the Clausius-Duhem inequality for the micropolar material expressed in terms of the rates of linearized kinematic variables as a guide to constitutive equation development, i.e.,

$$\boldsymbol{\sigma} : \dot{\bar{\boldsymbol{\varepsilon}}} + \mathbf{m} : \dot{\bar{\boldsymbol{\kappa}}} - \rho(\dot{\psi} + \eta\dot{T}) - \frac{1}{T} \mathbf{q} \cdot \nabla T \geq 0 \quad (53)$$

We assume a general form of the free energy that depends on the elastic strain, the elastic torsion-curvature, temperature, and a set of internal state variables,

i.e., $\psi = \hat{\psi}(\bar{\varepsilon}^e, \bar{\kappa}^e, T, \zeta^\alpha)$. Taking the chain rule expression for the time derivative of $\hat{\psi}$ along with the elastic-plastic decompositions for the strain and torsion-curvature and inserting into Eq. (53) and following the Coleman-Noll procedure, the following state laws are obtained:

$$\boldsymbol{\sigma} = \rho \frac{\partial \psi}{\partial \bar{\varepsilon}^e}, \quad \mathbf{m} = \rho \frac{\partial \psi}{\partial \bar{\kappa}^e}, \quad \eta = -\frac{\partial \psi}{\partial T} \quad (54)$$

along with the expression for the intrinsic dissipation, i.e.,

$$\Delta^{intr} = \boldsymbol{\sigma} : \dot{\bar{\varepsilon}}^p + \mathbf{m} : \dot{\bar{\kappa}}^p - \sum_{\alpha} r^{\alpha} * \dot{\zeta}^{\alpha} \geq 0, \quad r^{\alpha} = \rho \frac{\partial \psi}{\partial \zeta^{\alpha}} \quad (55)$$

Elastic Free Energy Function

The most general form of the elastic strain energy for a linearized micropolar continuum is given by the quadratic form viz (Eringen 1999):

$$\rho \psi_e = \frac{1}{2} \bar{\varepsilon}^e : \mathbb{C} : \bar{\varepsilon}^e + \frac{1}{2} \bar{\kappa}^e : \mathbb{D} : \bar{\kappa}^e + \bar{\varepsilon}^e : \mathbb{E} : \bar{\kappa}^e \quad (56)$$

where \mathbb{C} , \mathbb{D} , and \mathbb{E} are fourth-order tensors of elastic moduli. However, the coupling moduli \mathbb{E} are equal to zero for materials exhibiting point symmetry (Forest et al. 1997). Therefore, the stress and couple-stress constitutive equations for single crystals may be expressed as

$$\boldsymbol{\sigma} = \mathbb{C} : \bar{\varepsilon}^e, \quad \mathbf{m} = \mathbb{D} : \bar{\kappa}^e \quad (57)$$

For an elastically isotropic material, these expressions have the form

$$\begin{aligned} \boldsymbol{\sigma} &= \lambda \operatorname{tr}(\varepsilon^e) \mathbf{1} + 2\mu \varepsilon^e + 2\mu_c (\boldsymbol{\omega}^e - \bar{\boldsymbol{\Phi}}) \\ \mathbf{m} &= \alpha \operatorname{tr}(\bar{\kappa}^e) \mathbf{1} + 2\beta \operatorname{sym}(\bar{\kappa}^e) + 2\gamma \operatorname{skw}(\bar{\kappa}^e) \end{aligned} \quad (58)$$

where λ and μ are the usual Lamé's constants and μ_c , α , β , and γ are nonstandard and/or higher-order elastic moduli. The coupling modulus, μ_c , gives rise to the skew symmetric part of the Cauchy stress and the couple-stress moduli, α , β , and γ , can be interpreted as elastic length scales; when they are normalized, for example, with respect to the shear modulus the resulting quantities have units of length, e.g., $\ell_e = \sqrt{\beta/\mu}$.

Flow Rules

The development of dissipative constitutive equations follows the standard approach utilizing a flow potential and associative flow rules. As discussed in Forest and Sievert (2003), one can formulate the rules either in terms of a single or multiple criteria, i.e., either a single flow potential for both strain and torsion-curvature or independent flow potentials for the respective deformation variables. Mutli-criterion

theories are more general and require the specification of additional material parameters and should only be used if there is compelling need for this additional flexibility/complexity of the model. Representative multi-criterion micropolar single crystal plasticity models were presented by Forest et al. (2000) and Mayeur and McDowell (2013). The subsequent treatment of this section will focus on a single criterion framework since the numerical results presented later in the chapter were obtained using this type of model, and we have found it to suffice for most purposes. The plastic strain and torsion-curvature evolution equations are derived from a unified slip system flow potential, \mathcal{F}^α , i.e.,

$$\dot{\boldsymbol{\varepsilon}}^p = \sum_{\alpha} \dot{\lambda}^{\alpha} \frac{\partial \mathcal{F}^{\alpha}}{\partial \boldsymbol{\sigma}}, \quad \dot{\boldsymbol{\kappa}}^p = \sum_{\alpha} \dot{\lambda}^{\alpha} \frac{\partial \mathcal{F}^{\alpha}}{\partial \mathbf{m}} \quad (59)$$

Given the suggested kinematic forms for the plastic strain and torsion-curvature rates (Eqs. (49) and (51)), we propose a yield function, \mathcal{F}^α , of the form

$$\mathcal{F}^\alpha = \hat{\tau}^\alpha - (r_0 + r^\alpha) \leq 0 \quad (60)$$

Here, $\hat{\tau}^\alpha$ is an effective resolved shear stress and r_0 is the initial yield strength. The effective resolved shear stress is defined with respect to the projections of the force and couple stress tensors as

$$\hat{\tau}^\alpha = \sqrt{|\tau_{eff}^\alpha|^2 + |\pi_{\perp}^\alpha/L_{\perp}|^2 + |\pi_{\odot}^\alpha/L_{\odot}|^2}. \quad (61)$$

where τ_{eff}^α is the resolved shear stress, π_{\perp} is the resolved couple stress acting on edge GNDs, π_{\odot} is the resolved couple stress acting on screw GNDs, and L_{\perp} , L_{\odot} are normalizing length scales for edge and screw GNDs, respectively. The resolved shear and couple stresses are defined as

$$\tau_{eff}^\alpha = \boldsymbol{\sigma} : (\mathbf{s}^\alpha \otimes \mathbf{n}^\alpha) \quad (62)$$

$$\pi_{\perp}^\alpha = \mathbf{m} : (\mathbf{t}^\alpha \otimes \mathbf{s}^\alpha) \quad (63)$$

$$\pi_{\odot}^\alpha = \mathbf{m} : \left(\mathbf{s}^\alpha \otimes \mathbf{s}^\alpha - \frac{1}{2} \mathbf{1} \right) \quad (64)$$

The “*eff*” subscript has been applied to the resolved shear stress to emphasize that the driving force for slip has a contribution from the skew-symmetric part of the stress tensor and also to distinguish it from the classical resolved shear stress, which is computed using only the symmetric part of the Cauchy stress. The contribution of the skew-symmetric part of the Cauchy stress to τ_{eff}^α gives rise to gradient-dependent kinematic hardening which is elaborated upon further in section “[Relationship to Slip Gradient Theory](#).” The flow directions then follow as

$$\frac{\partial \mathcal{F}^\alpha}{\partial \boldsymbol{\sigma}} = \frac{\tau_{eff}^\alpha}{\hat{\tau}^\alpha} \mathbf{s}^\alpha \otimes \mathbf{n}^\alpha \quad (65)$$

$$\frac{\partial \mathcal{F}^\alpha}{\partial \mathbf{m}} = \frac{1}{L_\perp} \mathbf{t}^\alpha \otimes \mathbf{s}^\alpha \frac{\pi_\perp^\alpha / L_\perp}{\hat{\tau}^\alpha} + \frac{1}{L_\odot} \left(\mathbf{s}^\alpha \otimes \mathbf{s}^\alpha - \frac{1}{2} \mathbf{1} \right) \frac{\pi_\odot^\alpha / L_\odot}{\hat{\tau}^\alpha}. \quad (66)$$

Herein, we work within an elastic-viscoplastic setting and propose a power law expression for $\dot{\lambda}^\alpha$, i.e.,

$$\dot{\lambda}^\alpha = \dot{\lambda}_0 \left\langle \frac{\hat{\tau}^\alpha - (r_0 + r^\alpha)}{g^\alpha} \right\rangle^m \quad (67)$$

where $\dot{\lambda}_0$ is a reference effective deformation rate, g^α is a viscous drag stress, and m is an inverse rate sensitivity exponent. Inserting Eqs. (65), (66), and (67) into Eq. (59) and comparing with the expressions given in Eqs. (49) and (51), the expressions for the slip system deformation rates are obtained as

$$\dot{\gamma}^\alpha = \dot{\lambda}_0 \left\langle \frac{\hat{\tau}^\alpha - (r_0 + r^\alpha)}{g^\alpha} \right\rangle^m \frac{\tau_{eff}^\alpha}{\hat{\tau}^\alpha} \quad (68)$$

$$\dot{\varphi}_\perp^\alpha = \dot{\lambda}_0 \left\langle \frac{\hat{\tau}^\alpha - (r_0 + r^\alpha)}{g^\alpha} \right\rangle^m \frac{\pi_\perp^\alpha / L_\perp}{\hat{\tau}^\alpha} \quad (69)$$

$$\dot{\varphi}_\odot^\alpha = \dot{\lambda}_0 \left\langle \frac{\hat{\tau}^\alpha - (r_0 + r^\alpha)}{g^\alpha} \right\rangle^m \frac{\pi_\odot^\alpha / L_\odot}{\hat{\tau}^\alpha} \quad (70)$$

It is easily shown from Eqs. (68), (69), and (70) that the slip system plastic multiplier is related to the slip and curvature rates as

$$\dot{\lambda}^\alpha = \sqrt{|\dot{\gamma}^\alpha|^2 + |\dot{\varphi}_\perp^\alpha|^2 + |\dot{\varphi}_\odot^\alpha|^2} \quad (71)$$

Equation (71) reveals that $\dot{\lambda}^\alpha$ is an effective slip system deformation rate accounting for both slip and torsion-curvature deformation modes.

Internal State Variable Evolution

Consider the case where the evolution of the internal state variable associated with energetic isotropic hardening with the effective deformation rate, i.e.,

$$\dot{\zeta}^\alpha = \dot{\lambda}^\alpha \quad (72)$$

Further, a quadratic dependence of the free energy on ζ^α is assumed

$$\rho \psi_{in} = \frac{1}{2} \sum_\alpha \sum_\beta \mathbf{H}^{\alpha\beta} \zeta^\alpha \zeta^\beta, \quad \mathbf{H}^{\alpha\beta} = H_0 \quad (73)$$

where $H^{\alpha\beta}$ is the hardening matrix and $H_0 > 0$ is the hardening modulus. The energetic isotropic hardening stress is then given as

$$r^\alpha = \sum_{\beta} H^{\alpha\beta} \zeta^\beta = H_0 \sum_{\beta} \zeta^\beta \quad (74)$$

Given the preceding evolution equations, the intrinsic dissipation inequality may now be expressed in terms of slip system variables as

$$\Delta^{intr} = \sum_{\alpha} \left(\tau_{eff}^{\alpha} \dot{\gamma}^{\alpha} + \pi_{\perp}^{\alpha} \frac{\dot{\phi}_{\perp}^{\alpha}}{L_{\perp}} + \pi_{\odot}^{\alpha} \frac{\dot{\phi}_{\odot}^{\alpha}}{L_{\odot}} - r^{\alpha} \dot{\zeta}^{\alpha} \right) \geq 0 \quad (75)$$

which may be further simplified and expressed as

$$\Delta = \sum_{\alpha} (\hat{\tau}^{\alpha} - r^{\alpha}) \dot{\lambda}^{\alpha} \geq 0 \quad (76)$$

For nonzero values of $\dot{\lambda}^{\alpha}$, Eq. (67) can be inverted to obtain the expression for $\hat{\tau}^{\alpha}$ as

$$\hat{\tau}^{\alpha} = g^{\alpha} \left(\frac{\dot{\lambda}^{\alpha}}{\dot{\lambda}_0} \right)^{\frac{1}{m}} + r_0 + r^{\alpha} \quad (77)$$

Inserting this expression into Eq. (76) yields

$$\Delta^{intr} = \sum_{\alpha} \left[g^{\alpha} \left(\frac{\dot{\lambda}^{\alpha}}{\dot{\lambda}_0} \right)^{\frac{1}{m}} + r_0 \right] \dot{\lambda}^{\alpha} \geq 0 \quad (78)$$

Therefore, the dissipation inequality is unconditionally satisfied for this set of constitutive equations.

While the illustrative example given in this particular section yields a simple linear isotropic hardening behavior, it is straightforward and often preferable in practice to use nonlinear hardening laws; general examples of such extensions were previously given by Forest et al. (1997) and Mayeur et al. (2011). In the sequel, we present a dislocation density-based hardening framework (Mayeur and McDowell 2013) that has been compared to a number discrete dislocation dynamics boundary value problems; some of which will be presented as representative applications of the theory.

Dislocation Density-Based Strength Model

Previously, Mayeur et al. (2011) and Mayeur and McDowell (2013) have employed both single and mutli-criterion flow rules with a variety of hardening laws to simulate size-dependent behavior observed in 2D discrete dislocation dynamics

(DDD) simulations to varying degrees of success. The hardening model presented is the simplest version capable of reproducing the observed DDD behavior.

We take the statistically stored dislocation density on each slip system, ϱ_S^α , as our primary internal state variable to describe the hardening behavior. The slip system yield stress is defined in terms of a Taylor relation that is assumed to depend only on the SSD density, i.e.,

$$r^\alpha = r_0 + c_1 \mu b \sqrt{\sum_{\beta} h^{\alpha\beta} \varrho_S^\beta} \quad (79)$$

where c_1 is a constant, b is the Burgers vector, and $h^{\alpha\beta}$ is an interaction matrix. We do not use a generalized Taylor relation that directly includes a dependence on the GND density, which is a commonly employed assumption in other classes of generalized single crystal models. It was shown (Mayeur and McDowell 2013) that a hardening model based on the generalized Taylor relation leads to excessive and unrealistic hardening as compared to the DDD simulations. The SSD density evolves according to a Kocks-Mecking (Mecking and Kocks 1981) relation which represents a competition between storage and recovery mechanisms until the steady-state value of dislocation density is reached:

$$\dot{\varrho}_S^\alpha = \frac{1}{b} \left(\frac{1}{\Lambda^\alpha} - 2y_c \varrho_S^\alpha \right) \dot{\lambda}^\alpha \quad (80)$$

Here, Λ^α is the mean free path between dislocations, and y_c is the capture radius for dislocation annihilation. The mean free path is defined in terms of the SSD density, an average dislocation junction strength, K , and an interaction matrix, $a^{\alpha\beta}$ as

$$\Lambda^\alpha = \frac{K}{\sqrt{\sum_{\beta} a^{\alpha\beta} \varrho_S^\beta}} \quad (81)$$

Since the SSD evolution equation is defined with respect to $\dot{\lambda}^\alpha$ rather than $\dot{\gamma}^\alpha$, it naturally includes scale-dependence by virtue of $\dot{\lambda}^\alpha$ being an effective slip system deformation measure (see Eq. (71)). We found this type of scale-dependent isotropic hardening described the DDD results better than either a direct dependence of the Taylor stress and/or the mean free path on the GND density. We note that size effects of initial yield stress, r_0 , are not addressed by the hardening relations and may depend on the initial dislocation source distribution as well as obstacles and interfaces that impede slip.

Relationship to Slip Gradient Theory

In this section, we briefly discuss the relationship of micropolar crystal plasticity to slip gradient crystal plasticity theory in the form developed by Gurtin (2002, 2007).

A more in-depth comparison of the two model frameworks and simulation results has been given by Mayeur and McDowell (2014). Gurtin's theory is based on taking the slip system shears as continuum degrees of freedom and contains an additional balance equation – the so-called microforce balance – in addition to the classical force and angular momentum balances. The microforce balance is given as

$$\tau^\alpha + \nabla \cdot \xi^\alpha - q^\alpha = 0 \quad (82)$$

where τ^α is the classical resolved shear stress, ξ^α is a higher-order stress that is power-conjugate to the slip rate gradient, $\nabla \dot{\gamma}^\alpha$, and q^α is the stress power-conjugate to the slip rate, $\dot{\gamma}^\alpha$. In addition to Dirichlet micro-boundary conditions on the slip system shears, complementary micro-traction boundary conditions may also be prescribed in terms of the micro-traction Ξ^α , i.e.,

$$\Xi^\alpha = \xi^\alpha \cdot \mathbf{n} \quad (83)$$

The macroscopic and microscopic scales are coupled by the presence of the resolved shear stress, $\tau^\alpha = \mathbf{n}^\alpha \cdot \boldsymbol{\sigma} \cdot \mathbf{s}^\alpha$, in the microforce balance. The microforce balance is a partial differential equation that governs the evolution of the slip system shears and can be interpreted as a nonlocal yield condition.

Thermodynamics

Thermodynamically consistent constitutive equations have been developed by Gurtin using a purely mechanical form of the 2nd law, i.e.,

$$\Delta^{intr} = \boldsymbol{\sigma} : \dot{\varepsilon}^e + \sum_{\alpha} (q^\alpha \dot{\gamma}^\alpha + \xi^\alpha \cdot \nabla \dot{\gamma}^\alpha) - \dot{\psi} \geq 0 \quad (84)$$

The free energy is assumed to depend on the elastic strain and the set of slip system shear gradients, $\nabla \vec{\gamma} = \{\nabla \gamma^1, \dots, \nabla \gamma^N\}$, i.e.,

$$\psi = \hat{\psi}(\varepsilon^e, \nabla \vec{\gamma}) = \hat{\psi}_e(\varepsilon^e) + \hat{\psi}_{in}(\nabla \vec{\gamma}) \quad (85)$$

Expressing the time derivative of the free energy via the chain rule and inserting into Eq. (84) yields

$$\Delta^{intr} = \left(\boldsymbol{\sigma} - \frac{\partial \psi}{\partial \varepsilon^e} \right) : \dot{\varepsilon}^e + \sum_{\alpha} \left[q^\alpha \dot{\gamma}^\alpha + \left(\xi^\alpha - \frac{\partial \psi}{\partial \nabla \gamma^\alpha} \right) \cdot \nabla \dot{\gamma}^\alpha \right] \geq 0 \quad (86)$$

It is assumed that ξ^α can be decomposed additively into energetic and dissipative components such that

$$\xi^\alpha = \xi_{en}^\alpha + \xi_d^\alpha, \quad \xi_{en}^\alpha \doteq \frac{\partial \psi}{\partial \nabla \gamma^\alpha} \quad (87)$$

Following the Coleman-Gurtin (Coleman and Gurtin 1967) thermodynamic procedure, the state law for the Cauchy stress obtained is identical to that of Eq. (54)₁. Making use of Eq. (87), the reduced dissipation inequality is given as

$$\Delta^{intr} = \sum_{\alpha} (q^{\alpha} \dot{\gamma}^{\alpha} + \xi_d^{\alpha} \cdot \nabla \dot{\gamma}^{\alpha}) \geq 0 \quad (88)$$

Classical quadratic elastic free energy potentials are employed within the Gurtin-type theory since the treatment of the macroscopic forces is unaltered. Representative constitutive equations for the defect energy and dissipative stresses are presented below.

Energetic Constitutive Equations (Defect Energy)

Several variations of the defect energy have been proposed in the literature and can be classified as either recoverable or nonrecoverable. A recoverable defect energy depends on the current values of the slip gradients and vanishes when there are no slip gradients regardless of prior loading history, whereas a nonrecoverable energy does not vanish upon unloading. A typical recoverable defect energy is given as (Reddy 2011)

$$\psi_d = \frac{1}{2} g_0 \ell_{en}^2 \sum_{\alpha} \left[c_{\vdash} (\tilde{\varrho}_{\vdash}^{\alpha})^2 + c_{\odot} (\tilde{\varrho}_{\odot}^{\alpha})^2 \right] \quad (89)$$

where g_0 is the initial flow stress, ℓ_{en} is an energetic length scale, and c_{\vdash}/c_{\odot} are dimensionless constants defining the relative contributions of edge and screw GNDs, respectively. The edge and screw GND measures used in Eq. (89) are defined as

$$\tilde{\varrho}_{\vdash}^{\alpha} = -\nabla \gamma^{\alpha} \cdot \mathbf{s}^{\alpha}, \quad \tilde{\varrho}_{\odot}^{\alpha} = -\nabla \gamma^{\alpha} \cdot \mathbf{t}^{\alpha} \quad (90)$$

Note the dimensional and sign difference of the screw GND term with respect to the GND density definitions used in section “[Linearized Kinematics](#).” The energetic microcouple stress vector corresponding to ψ_d is given as

$$\xi_{en}^{\alpha} = -g_0 \ell_{en}^2 (c_{\vdash} \tilde{\varrho}_{\vdash}^{\alpha} \mathbf{s}^{\alpha} + c_{\odot} \tilde{\varrho}_{\odot}^{\alpha} \mathbf{t}^{\alpha}) \quad (91)$$

The microcouple stress vector lies in the slip plane and decomposes naturally into edge and screw components. Gurtin et al. (2007) view ξ_{en}^{α} as reflective of a net distributed Peach-Koehler force acting on the GNDs since its components are perpendicular to the respective dislocation line directions and have units of force per unit length.

Dissipative Constitutive Equations

The dissipative constitutive equations are introduced with the aid of an effective slip system deformation rate, d^{α} , defined as

$$d^\alpha = \sqrt{|\dot{\gamma}^\alpha|^2 + L_d^2 \|\nabla_{\tan}^\alpha \dot{\gamma}^\alpha\|^2} \quad (92)$$

where L_d is a dissipative length scale and the tangential gradient operator is defined as $\nabla_{\tan}^\alpha f = (\nabla f \cdot \mathbf{s}^\alpha) \mathbf{s}^\alpha + (\nabla f \cdot \mathbf{t}^\alpha) \mathbf{t}^\alpha$. The scalar dissipative microstress is posited to have the functional form

$$q^\alpha = g^\alpha \left(\frac{d^\alpha}{d_0} \right)^{\frac{1}{m}} \dot{\gamma}^\alpha \quad (93)$$

where g^α is a viscous drag stress, d_0 is the reference effective deformation rate, and m is the inverse rate-sensitivity exponent. The dissipative microcouple stress is introduced in an analogous manner, i.e.,

$$\xi_d^\alpha = g^\alpha L_d^2 \left(\frac{d^\alpha}{d_0} \right)^{\frac{1}{m}} \frac{\nabla_{\tan}^\alpha \dot{\gamma}^\alpha}{d^\alpha} \quad (94)$$

Inserting Eqs. (93) and (94) into Eq. (88) and making use of Eq. (92) yields

$$\Delta^{intr} = \sum_\alpha \tilde{\tau}^\alpha d^\alpha \geq 0, \quad \tilde{\tau}^\alpha := g^\alpha \left(\frac{d^\alpha}{d_0} \right)^{\frac{1}{m}}. \quad (95)$$

Here, we introduce the implicitly defined effective dissipative stress, $\tilde{\tau}^\alpha$, that is power-conjugate to the effective slip system deformation rate d^α . It is straightforward to show that $\tilde{\tau}^\alpha$ can also be expressed in terms of the dissipative microstresses, i.e.,

$$\tilde{\tau}^\alpha = \left[(q^\alpha)^2 + \|\xi_d^\alpha / L_d\|^2 \right]^{1/2}. \quad (96)$$

Inverted Flow Rule Versus Microforce Balance

An examination of the micropolar and slip gradient theories reveals many similarities between the two sets of governing equations. Central to both theories is the notion that the presence of GNDs gives rise to both energetic and dissipative contributions to scale-dependent mechanical behavior. In general, nonlocal strengthening effects are manifested in both the isotropic and kinematic hardening responses of the material. A key component of the slip gradient theory is the microforce balance, which couples the macroscopic and microscopic responses and represents a nonlocal flow rule governing the evolution of the slip system shears. Typically, the $\nabla \cdot \xi^\alpha$ term in the microforce balance induces kinematic hardening and the q^α term leads to isotropic hardening, although the exact nature of the hardening contributions depends on the specific constitutive forms employed. The noteworthy feature of second term in the microforce balance is that it is a function of second gradients of slip since $\xi^\alpha = f(\nabla_{\tan}^\alpha \dot{\gamma}^\alpha)$. The relationship between second gradients of slip

(GND gradients) and kinematic hardening is in accordance with other dislocation-based gradient single crystal plasticity theories (Yefimov et al. 2004a; Evers et al. 2004; Forest 2008).

A similar result is obtained for the micropolar theory by considering the inverted flow rule, which is obtained from the coaxiality relations $\tau_{eff}^\alpha/\hat{\tau}^\alpha = \dot{\gamma}^\alpha/\dot{\lambda}^\alpha$. Neglecting the isotropic energetic hardening terms in Eq. (77), the coaxiality relations may be rearranged and expressed as

$$\tau^\alpha + \tau_b^\alpha - g^\alpha \left(\frac{\dot{\lambda}^\alpha}{\dot{\lambda}_0} \right)^{\frac{1}{m}} \frac{\dot{\gamma}^\alpha}{\dot{\lambda}^\alpha} = 0. \quad (97)$$

Here, the resolved Cauchy stress, τ_{eff}^α , is explicitly written as two terms where the second term is a back stress, $\tau_b^\alpha := \frac{1}{2} \mathbf{t}^\alpha \cdot (\nabla \cdot \mathbf{m}^T)$, related to the projection of the skew symmetric part of the stress tensor. The back stress may be expressed as a function of the couple stress since $2 \text{skw}(\boldsymbol{\sigma}) = \boldsymbol{\epsilon} \cdot (\nabla \cdot \mathbf{m}^T)$. Thus, the back stress in the micropolar theory is a function of lattice torsion-curvature gradients. This is analogous to the dependence exhibited by the slip gradient theory since second gradients of slip (GND gradients) are directly connected to lattice torsion-curvature gradients via Nye's relations (see section "Linearized Kinematics"). In both theories, the higher-order balance law leads to the coupling of the micro and macro scales and to the natural inclusion of gradient-dependent kinematic hardening. Compare Eq. (97) to the microforce balance augmented with the constitutive equation for q^α given in Eq. (93)

$$\tau^\alpha + \nabla \cdot \boldsymbol{\xi}^\alpha - \tilde{g}^\alpha \left(\frac{d^\alpha}{d_0} \right)^{\frac{1}{m}} \frac{\dot{\gamma}^\alpha}{d^\alpha} = 0. \quad (98)$$

The similarities between the two expressions given in Eqs. (97) and (98) are obvious, although some key differences emerge when the constitutive equations for \mathbf{m} and $\boldsymbol{\xi}^\alpha$ are considered. To illustrate these differences, we compare the microforce balance and inverted micropolar flow rule augmented with simple constitutive equations for the 2D case (e.g., edge dislocations only). Inserting Eqs. (91) and (94) into Eq. (98), the following expression is obtained:

$$\begin{aligned} \tau^\alpha + g_0 \ell_{en}^2 c_t \nabla \cdot [(\nabla \gamma^\alpha \cdot \mathbf{s}^\alpha) \mathbf{s}^\alpha] + L_d^2 \nabla \cdot \left[g^\alpha \left(\frac{d^\alpha}{d_0} \right)^{\frac{1}{m}} \frac{(\nabla \dot{\gamma}^\alpha \cdot \mathbf{s}^\alpha) \mathbf{s}^\alpha}{d^\alpha} \right] \\ - g^\alpha \left(\frac{d^\alpha}{d_0} \right)^{\frac{1}{m}} \frac{\dot{\gamma}^\alpha}{d^\alpha} = 0. \end{aligned} \quad (99)$$

An analogous expression for the inverted micropolar flow rule is obtained by using a simplified version of Eq. (58)₂ where $\mathbf{m} = 2\mu \ell_e^2 \bar{\kappa}^e$ in Eq. (97), i.e.,

$$\tau^\alpha + \mu \ell_e^2 \mathbf{t}^\alpha \cdot (\nabla \cdot \bar{\boldsymbol{\kappa}}^T) - \frac{\mu \ell_e^2}{L_\perp} \mathbf{t}^\alpha \cdot \left(\nabla \cdot \sum_\beta \varphi_\perp^\beta \mathbf{s}^\beta \otimes \mathbf{t}^\beta \right) - g^\alpha \left(\frac{\dot{\lambda}^\alpha}{\dot{\lambda}_0} \right)^{\frac{1}{m}} \frac{\dot{\gamma}^\alpha}{\dot{\lambda}^\alpha} = 0. \quad (100)$$

Note that the energetic and dissipative length scales appear isolated in separate terms in Eq. (99), whereas the third term in Eq. (100) contains both the elastic and plastic length scales. Therefore, both energetic and dissipative gradient effects are eliminated when $\ell_e = 0$ in the micropolar theory; however, the dissipative gradient term is suppressed for $L_\perp \gg \ell_e^2$ irrespective of the value of ℓ_e . The separation of energetic and dissipative gradient hardening terms in the slip gradient theory is a direct consequence of the assumed additive decomposition of $\boldsymbol{\xi}$. Another, perhaps more subtle, difference between the two expressions is that the dissipative gradient term in the microforce balance is a function of slip rate gradients, while it is a function of total plastic curvature in the inverted micropolar flow rule. We remark that if the micropolar theory was cast in terms of an additive energetic/dissipative decomposition of the couple stress tensor rather than an additive elastic-plastic decomposition of the torsion-curvature tensor, the micropolar model would mirror the phenomenology of the slip gradient model in terms of a strict separation of energetic and dissipative gradient hardening effects, and the dissipative gradient hardening would be proportional to the rate of lattice torsion curvature.

Curvatures, Couple Stresses, and Gradient Terms

With the connection between the vectorial microcouple stress, $\boldsymbol{\xi}^\alpha$, and the second rank couple stress tensor, \mathbf{m} , established, additional analogies between the two models are explored. A comparison of Eqs. (90) and (52) reveals a relationship between the micropolar plastic torsion-curvature flow rates and the directional slip gradients, i.e.,

$$\dot{\varphi}_\perp^\alpha \leftrightarrow L_d \nabla \dot{\gamma}^\alpha \cdot \mathbf{s}^\alpha \quad , \quad \dot{\varphi}_\odot^\alpha \leftrightarrow L_d \nabla \dot{\gamma}^\alpha \cdot \mathbf{t}^\alpha \quad (101)$$

where \leftrightarrow is used to signify that the two terms have analogous roles in the respective theories. This association is further evident when the effective deformation rate in Eq. (92) is expressed in component form as

$$d^\alpha = \sqrt{|\dot{\gamma}^\alpha|^2 + |L_d \nabla \dot{\gamma}^\alpha \cdot \mathbf{s}^\alpha|^2 + |L_d \nabla \dot{\gamma}^\alpha \cdot \mathbf{t}^\alpha|^2}. \quad (102)$$

which establishes $d^\alpha \leftrightarrow \dot{\lambda}^\alpha$ (see Eq. (71)). If the energetic isotropic hardening term in Eq. (76) is neglected and compared to the analogous expression in Eq. (95), we see that the two dissipation inequalities have identical forms. Therefore, it is straightforward to show that the resolved couple stresses in the micropolar theory can be related to the dissipative microcouple stresses of the slip gradient theory by expressing the dissipation inequality in a form that parallels Eq. (75). First, the

dissipative constitutive equation for ξ_d^α given in Eq. (94) is written in component form as

$$\xi_d^\alpha = g^\alpha L_d^2 \left(\frac{d^\alpha}{d_0} \right)^{\frac{1}{m}} \frac{(\nabla \dot{\gamma}^\alpha \cdot \mathbf{s}^\alpha) \mathbf{s}^\alpha + (\nabla \dot{\gamma}^\alpha \cdot \mathbf{t}^\alpha) \mathbf{t}^\alpha}{d^\alpha} = \xi_{d\perp}^\alpha \mathbf{s}^\alpha + \xi_{d\circlearrowright}^\alpha \mathbf{t}^\alpha \quad (103)$$

where $\xi_{d\perp}^\alpha$ and $\xi_{d\circlearrowright}^\alpha$ are the edge and screw components of ξ_d^α . Inserting Eq. (103) into Eq. (88) and making use of the microforce balance, the dissipation inequality may be expressed as

$$\Delta^{intr} = \sum_\alpha \left[(\tau^\alpha + \nabla \cdot \xi^\alpha) \dot{\gamma}^\alpha + \frac{\xi_{d\perp}^\alpha}{L_d} (L_d \nabla \dot{\gamma}^\alpha \cdot \mathbf{s}^\alpha) + \frac{\xi_{d\circlearrowright}^\alpha}{L_d} (L_d \nabla \dot{\gamma}^\alpha \cdot \mathbf{t}^\alpha) \right] \geq 0. \quad (104)$$

A comparison of Eqs. (75) and (104) in conjunction with the relationships identified in Eq. (101) reveals that

$$\tau_{eff}^\alpha \leftrightarrow \tau^\alpha + \nabla \cdot \xi^\alpha \quad (105)$$

$$\pi_{\perp}^\alpha / L_{\perp} \leftrightarrow \xi_{d\perp}^\alpha / L_d \quad (106)$$

$$\pi_{\circlearrowright}^\alpha / L_{\circlearrowright} \leftrightarrow \xi_{d\circlearrowright}^\alpha / L_d \quad (107)$$

The association established in Eq. (105) was already apparent from a comparison of the inverted micropolar flow rule and the microforce balance, whereas the relationships established by Eqs. (106) and (107) were revealed by inserting the component form of ξ_d^α into the reduced dissipation inequality. Furthermore, we can rewrite the effective dissipative stress in the slip gradient theory using the component form of ξ_d^α as

$$\bar{\tau}^\alpha = \sqrt{|\tau^\alpha + \nabla \cdot \xi^\alpha|^2 + |\xi_{d\perp}^\alpha / L_d|^2 + |\xi_{d\circlearrowright}^\alpha / L_d|^2}. \quad (108)$$

This expression is consistent with Eq. (61) of the micropolar theory given the relationships established by Eqs. (105), (106), and (107).

As a final remark, we remind the reader that the micropolar theory differs from slip gradient theory by its adoption of the rotation of the microstructure director vectors in Eq. (30) as an independent constitutive prescription. This also simplifies the numerical implementation since slip gradients need not be computed or estimated. The remainder of section “[Micropolar Single Crystal Plasticity](#)” explores applications of this single crystal micropolar framework. The chapter dedicated to micromorphic crystal plasticity considers further extension of the micropolar theory to a micromorphic theory that includes distortion of the director vectors.

Applications: Comparison to 2D Discrete Dislocation Dynamics Simulations

This section presents a comparison of two-dimensional micropolar crystal plasticity and discrete dislocation dynamics simulations of single crystal thin films subjected to a variety of loading conditions. Specifically, we examine three widely studied boundary value problems: constrained simple shear, pure bending (single and double slip configurations), and simple shear of a particle reinforced composite. These problems are routinely used as benchmarks to evaluate the ability of generalized single crystal plasticity models to accurately capture size effects. A micropolar single crystal model with the hardening described by Eqs. (79), (80), and (81) was implemented in an Abaqus/Standard Version 6.7.1 (2007) user element subroutine (UEL) for the simulations presented below. The element is a four node quadrilateral and the integration is performed using a B-bar technique to prevent volumetric locking. Both the displacement and rotation fields are interpolated using standard bilinear shape functions.

The micropolar model has been independently calibrated for each boundary value problem by fitting to both average (e.g., stress-strain curves) and microscopic (e.g., dislocation density distributions) deformation behavior. It was previously demonstrated that fitting to multiple aspects of the deformation behavior is necessary to obtain a unique set of micropolar constitutive parameters (Mayeur et al. 2011). As a result of the 2D nature of the problems considered, the calibration procedure is simplified in the sense that there are only a few parameters to be determined. Essentially, there are four fitting parameters that are summarized by the set: $\mathbf{M} = \{r_0, K, \ell_e, L_\perp\}$. All of the other material parameters are either known as inputs (e.g., classical elastic parameters, Burgers vector magnitude) or can be assigned standard values based on experience from classical local crystal plasticity models. For example, the viscoplastic parameters $(\dot{\lambda}^\alpha, g^\alpha, m)$ cannot be explicitly determined due to lack of DDD results at multiple strain rates, so reasonable values for these parameters are prescribed and held fixed throughout the calibration process. Note that only a single elastic and plastic length must be determined. For plane strain problems, the couple stress constitutive relation given in Eq. (58)₂ reduces to

$$m_{3i} = 2\beta\bar{\kappa}_{3i}^e = 2\mu\ell_e^2\bar{\kappa}_{3i}^e, \quad i = 1, 2, \quad (109)$$

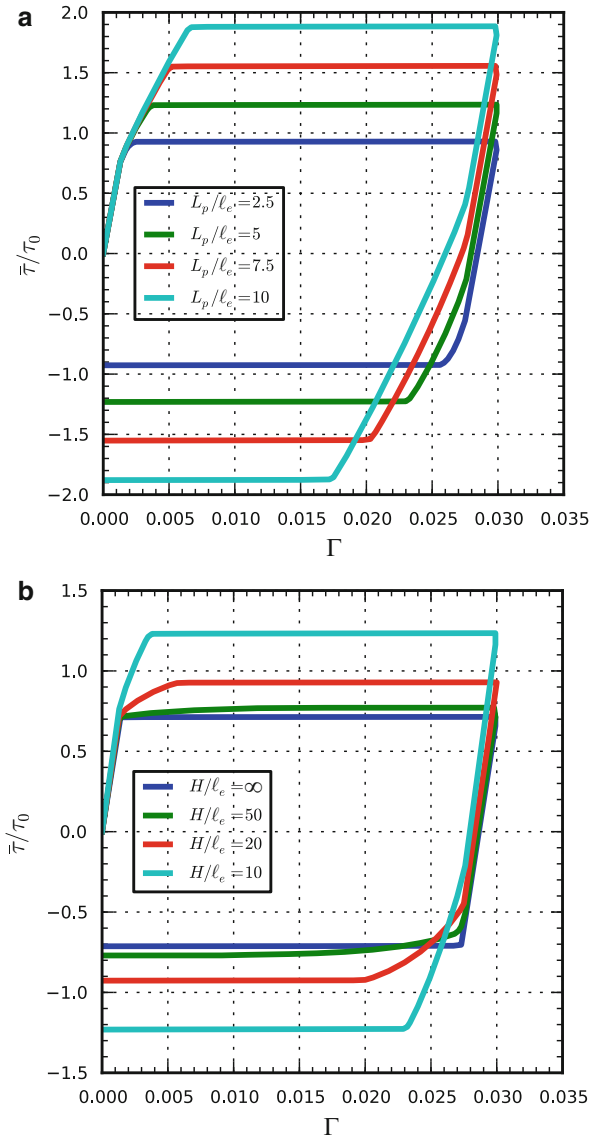
where it has been assumed, without loss of generality, that $\beta = \gamma$. Recall, the couple stress constitutive parameter, β , is related to an elastic length scale parameter, $\ell_e = \sqrt{\beta/\mu}$. As discussed by Forest (2008), the nonclassical elastic constant, μ_c , Eq. (58)₁, is not a free fitting parameter in micropolar single crystal plasticity. Rather, it serves as an internal penalty constraint forcing the lattice rotations to coincide with the rotational part of the elastic distortion. Because of this constraint, the micropolar torsion-curvature is identified as the lattice torsion-curvature thereby making the connection between GNDs and the scale effects predicted by the model.

Since screw GNDs do not contribute to plastic torsion-curvature evolution in the 2D boundary value problems, there is a single plastic length scale parameter, L_{\perp} , to be determined.

Here, we briefly outline how the length scale parameters affect material response a general strategy for calibrating them to the DDD simulations. The elastic length scale parameter, ℓ_e , is related to the initial scale-dependent kinematic hardening modulus and the ratio L_{\perp}/ℓ_e dictates the saturation rate of gradient hardening. Larger L_{\perp}/ℓ_e ratios are associated with slower transients such that when $L_{\perp} \gg \ell_e$, the strengthening effect reduces to that of linear gradient kinematic hardening (Mayeur et al. 2011). The elastic length scale should also be small enough so as not to induce size-dependent effects in the elastic deformation regime. The plastic length scale, L_{\perp} , determines the magnitude of the GND density distributions; however, its influence is not independent of the prescribed value of ℓ_e . As shown in Mayeur et al. (2011), it is the ratio L_{\perp}/ℓ_e that dictates the maximum value of the local GND density fields. If ℓ_e is varied while holding L_{\perp}/ℓ_e and all other model parameters constant (including specimen dimensions), the resulting GND density field will be essentially unchanged. To illustrate how the length scale parameters influence the average stress strain response, we show two plots in Fig. 1 where the elastic length scale is varied while keeping the ratio L_{\perp}/ℓ_e fixed (here $L_p = L_{\perp}$) and another with fixed elastic length scale and varying L_{\perp}/ℓ_e for the problem of constrained simple shear (Mayeur and McDowell 2014).

The parameter r_0 is used to fit to the initial yield strength rather than a prescribed initial dislocation density since the DDD simulations are assumed to be initially dislocation free. However, in the planar double slip simulations, a negligibly small value of initial SSD density ($10^{-6} \mu\text{m}^{-2}$) is specified for each slip system so that SSD evolution is nonzero. This approach is taken in lieu of introducing a nucleation term in the SSD density evolution equation. It is worth mentioning that r_0 should not necessarily be treated as a fixed material constant due to the statistical variations in the source strength and spatial distribution in the discrete dislocation simulations, as well as the mean free path of initial obstacles or impenetrable interfaces. The initial yield point in the discrete simulations will depend significantly on the availability of weak sources in highly stressed regions. Given a fixed value for the initial slip system yield strength, approximate upper bounds can be established for ℓ_e and K by isolating the effects of the two distinct material strengthening mechanisms. For example, given a target stress-strain curve, the maximum value of ℓ_e is determined by assuming that all of the material hardening is due to the gradient-induced back stresses ($1/K = 0$), whereas the maximum value of K is determined by assuming that all of the material hardening is due to slip threshold hardening ($\ell_e = 0$). In general, both mechanisms will contribute to the material strengthening and the actual values will fall below the upper bounds. As discussed in Mayeur and McDowell (2013), either the unloading behavior or some other attribute of the local deformation field (e.g., shear strain distributions) must be used in order to differentiate between the relative hardening contributions. A list of micropolar constitutive parameters used in all of the simulations is given in Table 1.

Fig. 1 Average shear stress-strain response for constrained simple shear with (a) different values of L_p/ℓ_e with $\ell_e = 100$ nm and (b) different elastic length scale parameters with $L_p/\ell_e = 5$. Slip threshold hardening is suppressed ($r^\alpha = r_0$)



Constrained Shear of Thin Films

Here, we compare results of the micropolar model to DDD results for a constrained thin film subjected to simple shear originally presented by Shu et al. (2001). Related studies were also conducted by Yefimov and van der Giessen (2005b) and Limkumnerd and van der Giessen (2008). A more detailed exposition of the micropolar results are contained in Mayeur and McDowell (2013). The specimen

Table 1 Summary of constitutive model parameters used in the micropolar single crystal simulations

Parameter	Symbol		Magnitude			Unit
–	–	Constrained shear	Bending SS	Bending DS	Composite	–
Shear modulus	μ	26.3	26.3	26.3	26.3	GPa
Poisson's ratio	ν	0.33	0.33	0.33	0.33	–
Couple modulus	μ_c	263	263	263	263	GPa
Elastic length scale	ℓ_e	10, 15	125, 125	300, 600	125	nm
Plastic length scale	L_{\perp}	45	562.5, 250	750, 700	125	nm
Reference threshold stress	r_0	12.78	10	10	13, 21, 30	MPa
Threshold stress coefficient	c_1	0.5	N/A	0.5	N/A	–
Burgers vector magnitude	b	0.25	0.25	0.25	nm	–
Hardening matrix coefficients	$h^{\alpha\beta}$	$\delta^{\alpha\beta}$	N/A	1.0	N/A	–
Initial SSD density	ρ_{s0}	10^{-6}	N/A	10^{-6}	N/A	μm^{-2}
Dislocation interaction coefficients	$\alpha^{\alpha\beta}$	$\delta^{\alpha\beta}$	N/A	1.0	N/A	–
Dislocation segment length constant	K	16.67, 18.18	N/A	160, 26	N/A	–
Dislocation capture radius	y_c	0	N/A	1.5	N/A	nm
Reference deformation rate	$\dot{\lambda}_0$	10^{-3}	10^{-3}	10^{-3}	10^{-3}	s^{-1}
Drag stress	g	5	5	5	5	MPa
Inverse rate sensitivity exponent	m	20	20	20	20	–

geometry is film oriented for symmetric slip with thickness, H , in the x_2 -direction as shown in Fig. 2. The slip systems are oriented at $\pm 30^\circ$ with respect to the x_2 -direction. The top and bottom surfaces are modeled as rigid dislocation barriers (impenetrable), the bottom surface is fully constrained against displacement, and a uniform horizontal displacement is applied to the top surface while the vertical displacement is constrained. The load is applied under displacement control up to an

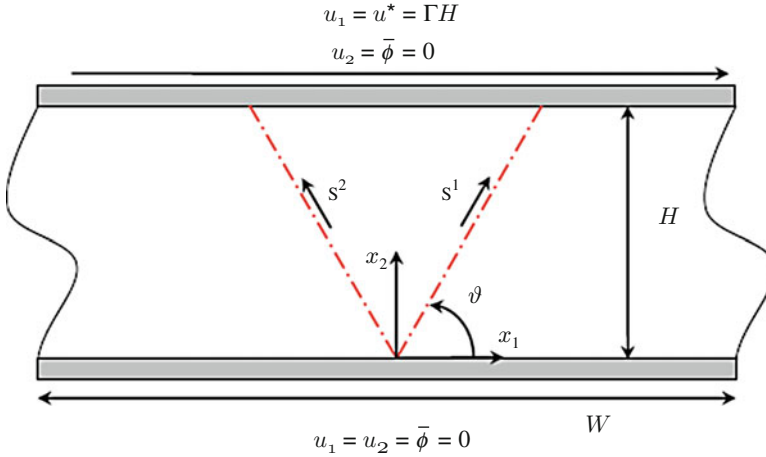


Fig. 2 Geometry schematic and boundary conditions for the constrained shear initial-boundary value problem

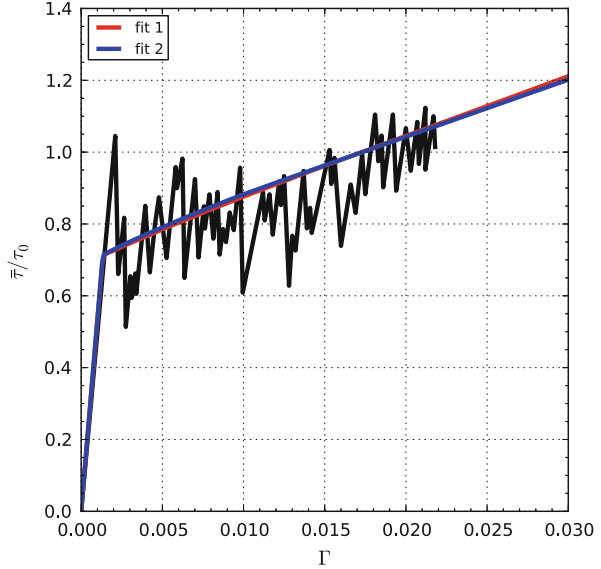
average strain of $\Gamma = 0.03$. The discrete dislocation problem was modeled as a unit cell of width, W , and thickness, H , and was spatially discretized with uniformly sized quadrilateral finite elements with an element size of $h_e = W/30$. The material was modeled as having zero initial dislocation density and the sources were distributed randomly throughout the spatial domain. Individual dislocation source strengths were determined by randomly sampling from a Gaussian distribution with a mean nucleation strength, $\bar{\tau}_{nuc} = 50$ MPa, and a standard deviation of $0.2\bar{\tau}_{nuc}$.

The displacement boundary conditions were applied at the top surface consistent with a constant strain rate of 10^3 s^{-1} . The constitutive parameters used in the simulations are representative of an aluminum single crystal.

$$\begin{aligned}
 u_1(x_1, 0, t) &= u_2(x_1, 0, t) = \bar{\phi}(x_1, 0, t) = 0 \\
 u_1(x_1, H, t) &= \Gamma(x_1, H, t)H, \quad u_2(x_1, H, t) = \bar{\phi}(x_1, H, t) = 0 \\
 u_1\left(\frac{W}{2}, x_2, t\right) &= u_1\left(-\frac{W}{2}, x_2, t\right), \quad u_2\left(\frac{W}{2}, x_2, t\right) = u_2\left(-\frac{W}{2}, x_2, t\right) \\
 \bar{\phi}\left(\frac{W}{2}, x_2, t\right) &= \bar{\phi}\left(-\frac{W}{2}, x_2, t\right)
 \end{aligned} \tag{110}$$

We only used the data for the $1 \mu\text{m}$ thick film in the calibration process since this is the only thickness for which all relevant deformation fields have been reported. Consistent with earlier nonlocal single crystal plasticity simulations of this problem, the isotropic hardening response is assumed to be linear (Shu et al. 2001; Bittencourt et al. 2003). Due to the symmetry of the problem and since it is impossible to differentiate between self and latent hardening effects from the available DDD results, we assume that $h^{\alpha\beta} = a^{\alpha\beta} = \delta^{\alpha\beta}$. Results are presented for two sets of constitutive parameters. As shown in Fig. 3, the stress-strain curves are nearly identical for both fits and are in good agreement with the discrete dislocation results. Fit 1 uses slightly

Fig. 3 Comparison of the average shear stress-strain response for $H = 1 \mu\text{m}$. Results for two parameter sets are shown to illustrate how differences in the local fields can vary for an identical average response. The discrete dislocation results are from Shu et al. (2001) and are given by the solid black line



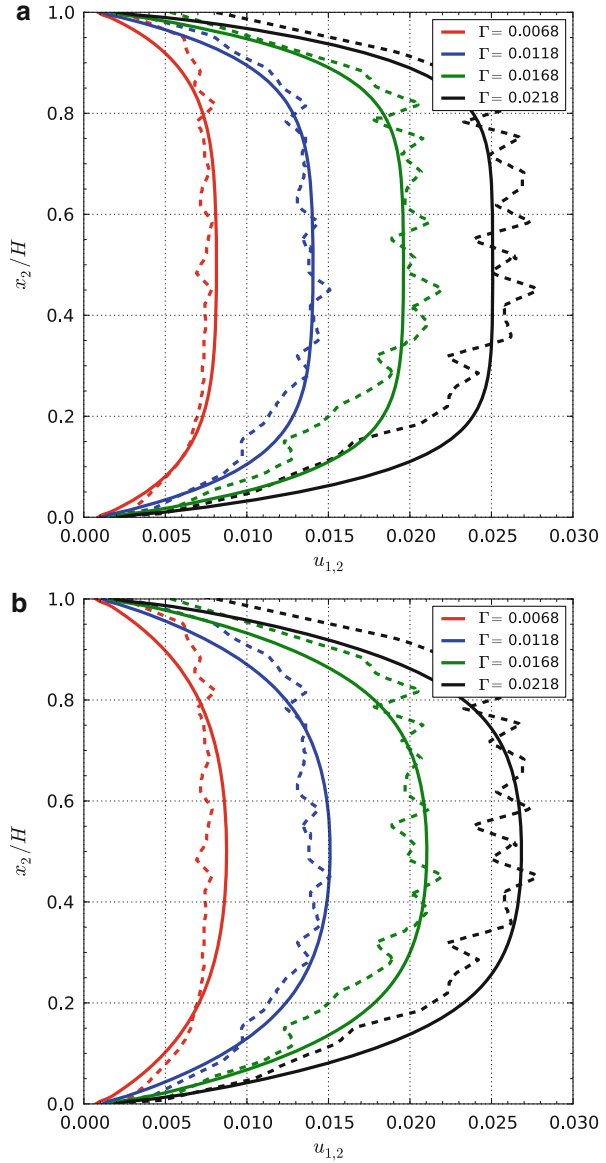
more threshold hardening ($K_1 = 16.67$ vs. $K_2 = 18.18$), whereas fit 2 has a larger contribution from gradient kinematic hardening ($\ell_{e2} = 15$ nm vs. $\ell_{e1} = 10$ nm). The differences in the relative proportions of isotropic versus gradient kinematic hardening are evident in the shear strain distributions shown in Fig. 4. The profiles for fit 1 have a blunted shape as compared to the rounded morphology observed for fit 2, which is consistent with the general observation that larger elastic length scales result in parabolic shear strain distributions. The signed GND density distributions for both fits are shown in Fig. 5a. The maximum GND densities at the boundary are marginally overpredicted and display steeper gradients in the near boundary regions than the discrete dislocation results, but compare favorably overall.

Pure Bending of Thin Films: Single Slip

The discrete dislocation results of the single slip bending configuration presented in this section were obtained by Yefimov et al. (2004b) and the micropolar results were given by Mayeur and McDowell (2011). The initial-boundary value problem is a thin film of width, W , and thickness, H , subjected to pure bending in a state of plane strain as sketched in Fig. 6. Considering a coordinate system attached to the midpoint of the film, the deformation is defined by the edge rotation angle, Θ , and is prescribed through a linear variation of the x_1 displacement component as a function of distance from the neutral axis:

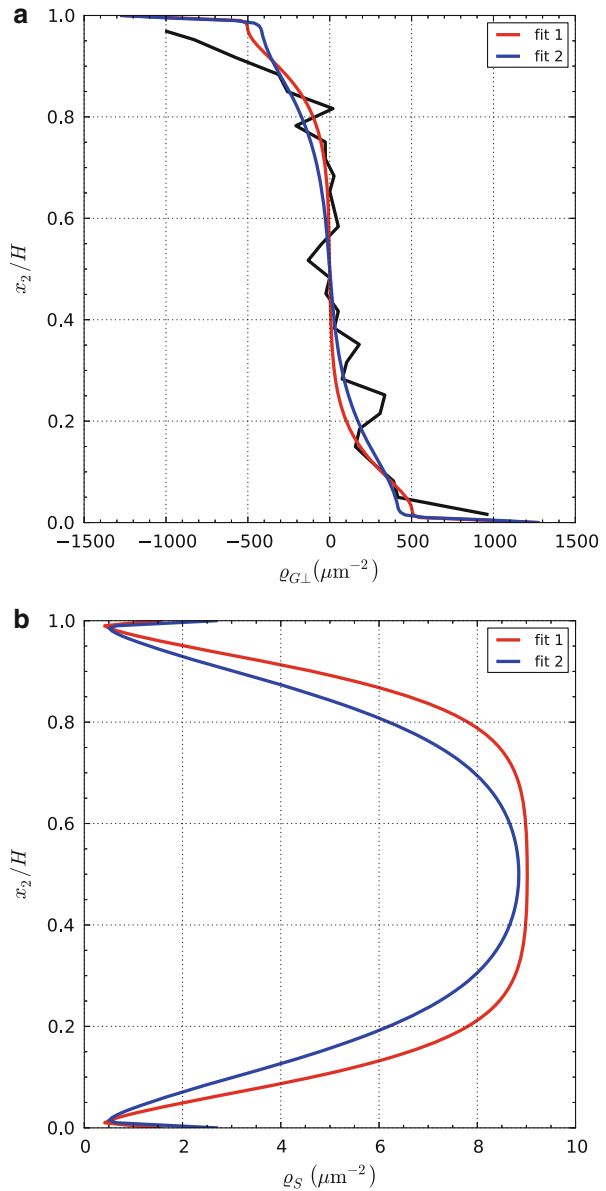
$$u_1 \left(\pm \frac{W}{2}, x_2, t \right) = \pm \Theta(t) x_2. \quad (111)$$

Fig. 4 Comparison of shear strain distributions at different levels of applied strain for $H = 1 \mu\text{m}$: (a) fit 1 and (b) fit 2. The discrete dislocation results are from Shu et al. (2001) and are given by the dashed lines



The displacements are applied consistent with a constant average rotation rate, 10^{-3} s^{-1} , until a final rotation angle of $\Theta = 0.02$ is reached. The top and bottom surfaces of the beam are traction-free. As shown in Fig. 6, a single slip system is oriented at an angle ϑ with respect to the x_1 -axis, and slip is constrained to occur within region demarcated by the internal solid black lines. This restriction has been imposed in the discrete dislocation simulations in order to avoid the complication of

Fig. 5 Comparison of discrete dislocation and micropolar crystal plasticity dislocation density distributions at $\Gamma = 0.0168$ for $H = 1 \mu\text{m}$: (a) signed GND density and (b) SSD density. The discrete dislocation result is from Shu et al. (2001) and are given by the solid black line



having dislocations exit the crystal through the lateral faces where the displacement boundary conditions are prescribed. Two different film thicknesses are considered for the single slip configuration, $H = 2$ and $4 \mu\text{m}$, respectively, with a fixed width-to-thickness ratio of $W/H = 3$. Slip system orientations of 30° and 60° are studied and will be referred to using the shorthand notation ϑ_{30} and ϑ_{60} in the following.

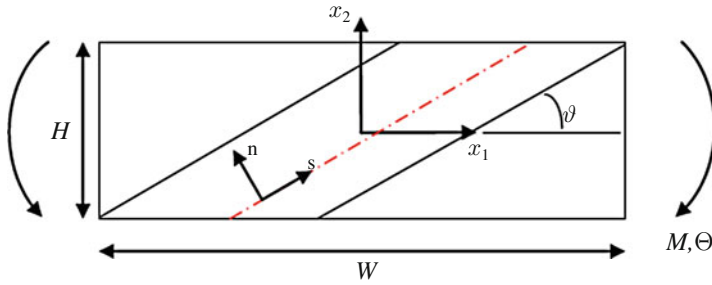


Fig. 6 Schematic of the geometry and slip system configuration for the single slip bending initial-boundary value problem

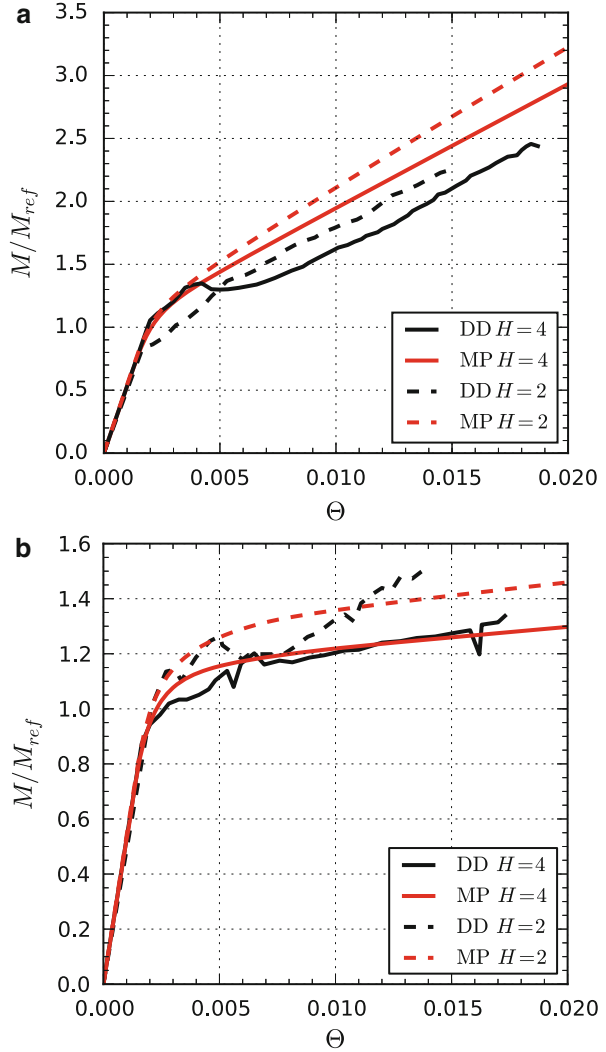
The average loading response is quantified by the bending moment, M , work-conjugate to Θ which is given by

$$M = \int_{-H/2}^{H/2} \sigma_{11} \left(\pm \frac{W}{2}, x_2 \right) x_2 dx_2. \quad (112)$$

A thickness-independent measure of the average loading response is given by the normalized bending moment, M/M_{ref} , with the reference bending moment defined as: $M_{ref} = \frac{2}{3} \bar{\tau}_{nuc} \left(\frac{H}{2} \right)^2$. M_{ref} corresponds to the moment calculated from an assumed linear stress distribution over height of the beam with peak values of $\pm \bar{\tau}_{nuc}$ at the free surfaces. The mean critical nucleation stress, $\bar{\tau}_{nuc}$, is taken as 50 MPa in the discrete dislocation simulations. The micropolar finite element meshes, consistent with the DDD simulations, employ a uniform grid of bilinear quadrilaterals: 66×38 for ϑ_{30} and 155×30 for ϑ_{60} .

The normalized moment-rotation plots for both slip system orientations and thicknesses are plotted in Fig. 7, and they are in reasonably good agreement with the discrete dislocation results with respect to both the orientation and scale-dependence. Except for the $2 \mu\text{m}$ thick film for ϑ_{30} , the results are in good quantitative agreement with the initial yield strengths and nominal hardening rates predicted by the discrete dislocation model. The DDD results show that the yield strength for the $2 \mu\text{m}$ thick film for ϑ_{30} is lower than that of the $4 \mu\text{m}$ film in contradiction to an expected “smaller is stronger” trend, which underscores the stochastic nature of the initial flow stress obtained from DDD simulations. The discrete simulations display an approximately linear hardening rate that increases with decreasing film thickness, and the micropolar model shows similar trends although the rate is somewhat underestimated for the $2 \mu\text{m}$ thick film with ϑ_{60} . There is a substantially higher hardening rate for ϑ_{30} as compared to ϑ_{60} as shown in Fig. 7; this difference is primarily due to the way the boundary value problem is constructed with distinct elastic and plastic zones and is not a consequence of the local hardening behavior. Recall that the films are modeled as composite elastic-plastic materials with plastic deformation restricted to the interior region

Fig. 7 Comparison of the discrete dislocation (DD) and micropolar crystal plasticity (MP) normalized moment-rotation response for the single slip bending configuration (a) $\vartheta = 30^\circ$ and (b) $\vartheta = 60^\circ$. The discrete dislocation results are from Yefimov et al. (2004b)



bounded by the solid lines parallel to the slip direction (see Fig. 6) for the sake of convenience in the DDD simulations. Therefore, the plastic zone size for ϑ_{30} is much smaller. The significantly higher apparent hardening rates for ϑ_{30} as compared to ϑ_{60} are essentially due to an increased volume fraction of the elastic phase and not dislocation hardening. In fact, a local crystal plasticity model with an elastic-perfectly plastic slip system level response would yield an apparent hardening rate comparable to, albeit lower than that shown in Fig. 7a. Of course, there is a component of the apparent hardening rate differences for the two orientations due to the relative misalignment of the slip and axial strain directions, and this is the portion associated with variations in GND distributions between the two cases.

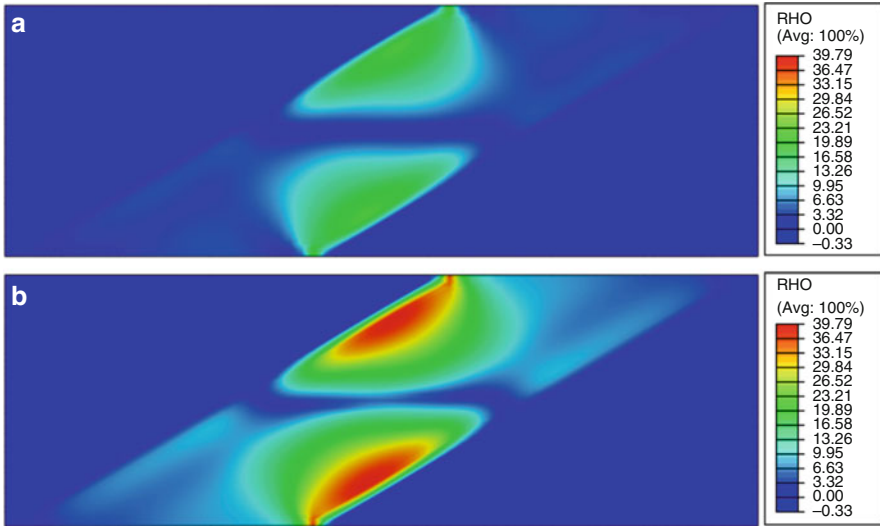


Fig. 8 Dislocation density distributions predicted by the micropolar crystal plasticity simulations at $\Theta = 0.015$ for $\vartheta = 30^\circ$: (a) $H = 4 \mu\text{m}$ (b) $H = 2 \mu\text{m}$

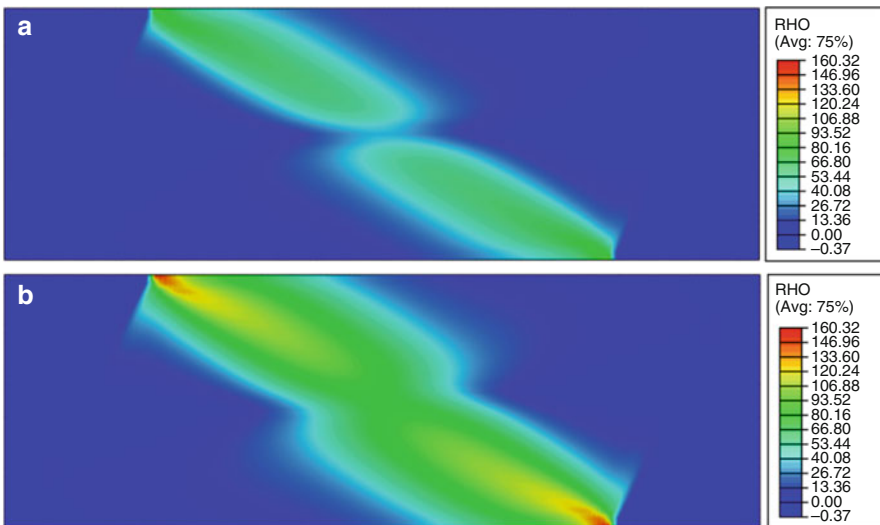
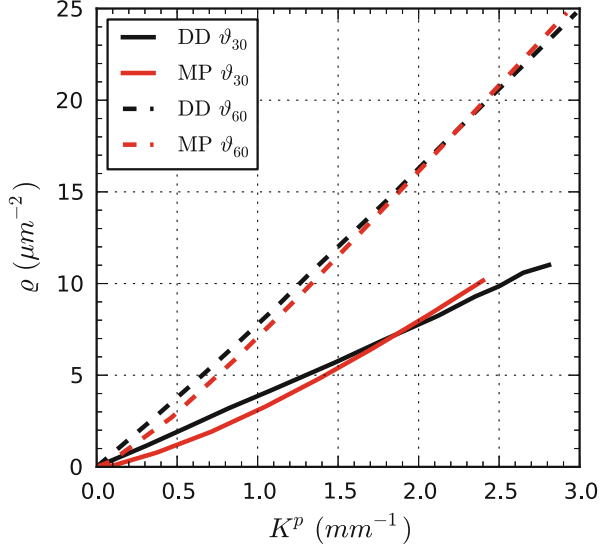


Fig. 9 Dislocation density distributions predicted by the micropolar crystal plasticity simulations at $\Theta = 0.015$ for $\vartheta = 60^\circ$: (a) $H = 4 \mu\text{m}$ (b) $H = 2 \mu\text{m}$

Dislocation density contour plots for both film thicknesses are shown in Figs. 8 and 9 for ϑ_{30} and ϑ_{60} , respectively, with $\Theta = 0.015$. In the ϑ_{30} film, the maximum dislocation densities are $39.8 \mu\text{m}^{-2}$ and $22.6 \mu\text{m}^{-2}$, respectively, for 2 and 4 μm thick films. Dislocation-free zones are clearly observed along the neutral

Fig. 10 Comparison of the discrete dislocation (DD) and micropolar crystal plasticity (MP) dislocation density evolution as a function of average plastic curvature for both orientations of the single slip configuration. The dashed curves are for ϑ_{60} . The discrete dislocation results are from Yefimov et al. (2004b)



axis for both film thicknesses, where the thickness of the dislocation-free region is approximately 3–4 times larger for the 4 μm thick film. As compared to the ϑ_{30} film, the dislocation density distributions for ϑ_{60} are markedly different. The morphology of the distribution can be characterized as having ellipsoidal-shaped lobes originating at the free surface near the corner of the elastic-plastic interface and extending perpendicular to elastic-plastic interface toward the neutral axis. The maximum local dislocation densities, 160 μm^{-2} ($H = 2 \mu\text{m}$) and 98.3 μm^{-2} ($H = 4 \mu\text{m}$), are significantly higher than in the ϑ_{30} film. The contour plots exhibit rather high dislocation density at the free surfaces that are generated in response to the strong rotational gradients that arise due to the compliance mismatch at the elastic-plastic interface. This is in contrast to the DDD simulations which exhibit dislocation-free zones at the free surfaces that are thought to be the result of an image force effect. It is possible that an image force effect might be imposed within the micropolar framework through an appropriately specified higher-order traction along the free surfaces; however, this avenue has yet to be pursued.

In Fig. 10, the total dislocation density computed over the entire volume is plotted as a function of the imposed deformation for both orientations with $H = 4 \mu\text{m}$. It is shown that the micropolar model accurately captures the evolution as predicted by the DDD simulations. The dislocation density is computed by volume averaging the centroidal element values over the FE mesh, and the average plastic curvature, K^p , is calculated according to

$$K^p = \frac{2\Theta}{W} - \frac{M}{EI}. \quad (113)$$

Here, EI is the in-plane bending stiffness defining the elastic curvature. The total GND density required to accommodate an imposed bending angle, Θ , can be calculated according to Ashby (1970) in terms of K^p as

$$\hat{\rho} = \frac{K^p}{b_1}, \quad (114)$$

where b_1 is the magnitude of the x_1 -component of the Burgers vector. The plot shows that the dislocation density increases in an approximately linear fashion with respect to the plastic curvature for both sets of simulations. Also, as expected from Eq. (114), the dislocation density at a given level of plastic curvature is higher for the ϑ_{60} film. Since the x_1 -component of the Burgers vector is smaller for this orientation, more dislocations are needed to accommodate the imposed strain gradient.

Pure Bending of Thin Films: Double Slip

Next we study pure bending of thin films with a double slip system configuration as shown in Fig. 11. The DDD simulation results presented were originally reported by Yefimov and van der Giessen (2005a) and an analysis using micropolar crystal plasticity was presented by Mayeur (2010). As in the single slip case, two orientations are considered, ϑ_{30} and ϑ_{60} where the orientation angle is also defined as the angle between the x_1 -axis and the slip direction for slip system 1. For each orientation, the second slip system is symmetrically aligned with respect to the x_2 -axis. The films are partitioned into elastic and plastic phases as before and in accordance with the discrete dislocation simulations. The elastic zones are the triangular regions located at the top, bottom, left, and right ends of the film as shown in Fig. 11, and each slip system is only active in the slice of material parallel to the slip direction. Thus, the plastic zone can be divided into five regions: four single slip regions (two for each slip system) that are the outermost diagonal regions and a diamond-shaped double slip region located at the center of the film and is demarcated by the blue dotted lines in the schematic. The boundary conditions and

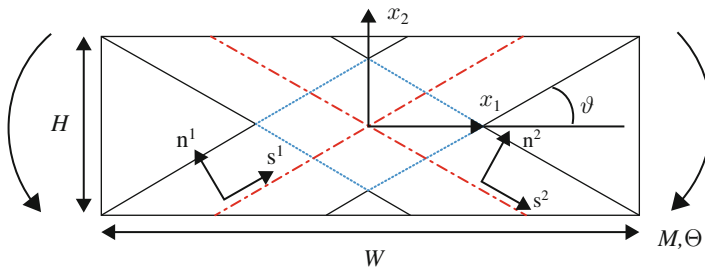
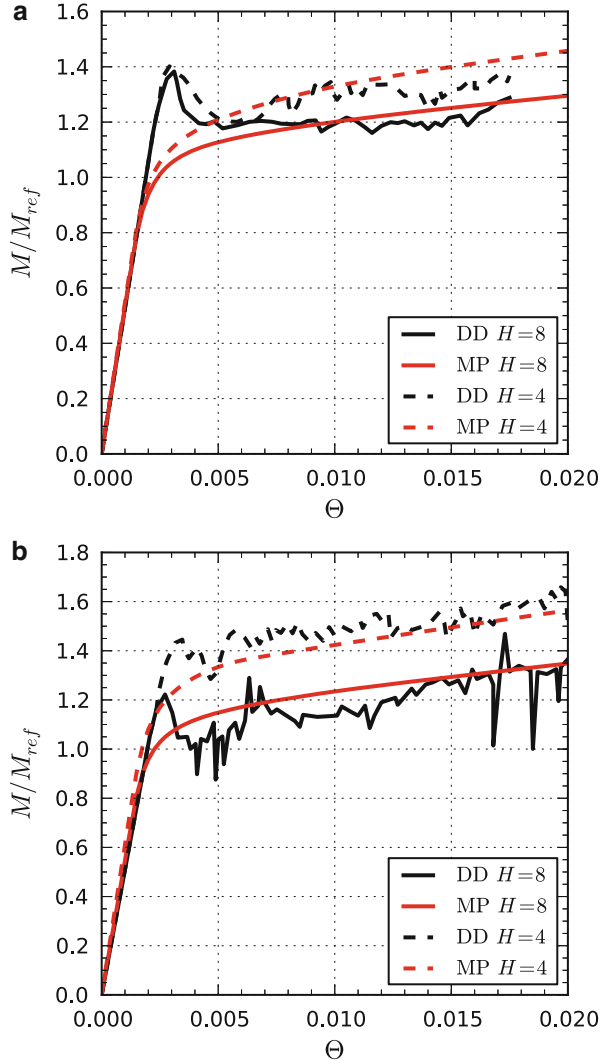


Fig. 11 Schematic of the geometry and slip system configuration for the double slip bending initial-boundary value problem

Fig. 12 Comparison of the discrete dislocation (DD) and micropolar crystal plasticity (MP) normalized moment-rotation response for the double slip bending configuration (a) $\vartheta = 30^\circ$ and (b) $\vartheta = 60^\circ$. The discrete dislocation results are from Yefimov and van der Giessen (2005a)



FE discretizations are the same as for the single slip configuration. Simulations are carried out for film thicknesses of $4 \mu\text{m}$ and $8 \mu\text{m}$ with fixed aspect ratio $W/H = 3$.

The normalized moment-rotation responses for both film thicknesses and orientations are plotted in Fig. 12 against the discrete dislocation results, and the results compare favorably. The response is similar for both orientations unlike the single slip configuration, where the hardening rate was much higher for the ϑ_{30} films due to larger effective film thickness resulting from the dominant influence of the elastic regions. However, the behavior for the ϑ_{60} films has a stronger scale-dependence as would be expected given that more GNDs are necessary to accommodate the

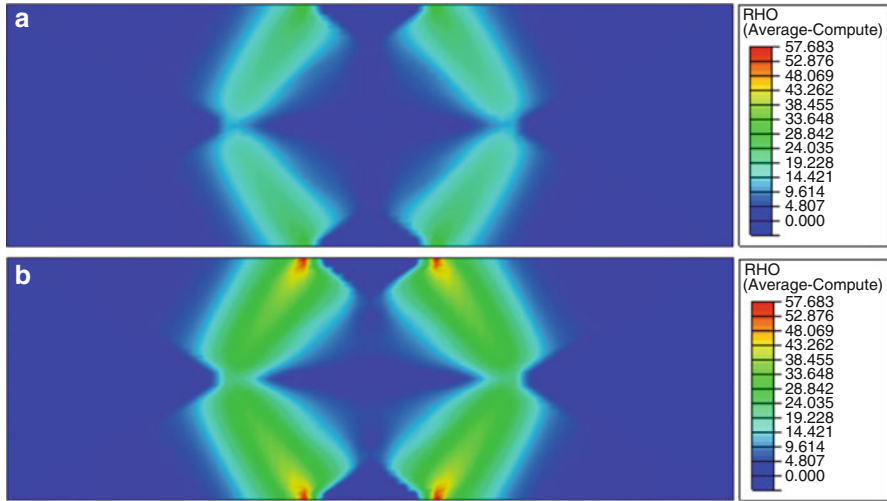


Fig. 13 Dislocation density distributions predicted by the micropolar crystal plasticity simulations at $\Theta = 0.02$ for $\vartheta = 30^\circ$: (a) $H = 8 \mu\text{m}$ (b) $H = 4 \mu\text{m}$

strain gradients for the ϑ_{60} film. In contrast to the single slip simulations, we found that different elastic length scales and dislocation multiplication constants, K , are required for each orientation to obtain a good match with the DDD results.

Figure 13 shows the total dislocation density contour plots for the ϑ_{30} films at $\Theta = 0.02$, and we note that the magnitude of the total dislocation density field is approximately three orders of magnitude larger than that of the SSD density field (not shown), which is characteristic of confined micro-scale dislocation plasticity. The maximum dislocation density is located at the free surfaces near the elastic-plastic interfaces for both film thicknesses, with peak values of $57.7 \mu\text{m}^{-2}$ and $32.3 \mu\text{m}^{-2}$ for the $4 \mu\text{m}$ and $8 \mu\text{m}$ thick films, respectively. In the case of the $8 \mu\text{m}$ thick film, there is dislocation-limited region (not dislocation-free) adjacent to the neutral axis separating the regions of higher dislocation density, whereas the dislocation distributions are continuous across the neutral axis for the $4 \mu\text{m}$ thick film. This is in contrast to the dislocation density fields for the single slip configuration (see Fig. 8) which exhibit a clear dislocation starved zone adjacent to the neutral axis. It is interesting that the maximum dislocation density values occur in the single slip regions and that the double slip region has a significantly lower density.

Figure 14 shows the total dislocation density contour plots for the ϑ_{60} films at $\Theta = 0.02$. The geometrical configuration of the elastic and plastic phases for the ϑ_{60} oriented crystal is such that there is no centrally located elastic zone and the majority of the plastic phase is a double slip region. Therefore, the dislocation density fields are continuous and smooth, in contrast to the ϑ_{30} orientation where the dislocation density field has a checkered type of pattern. The local maximum in the dislocation

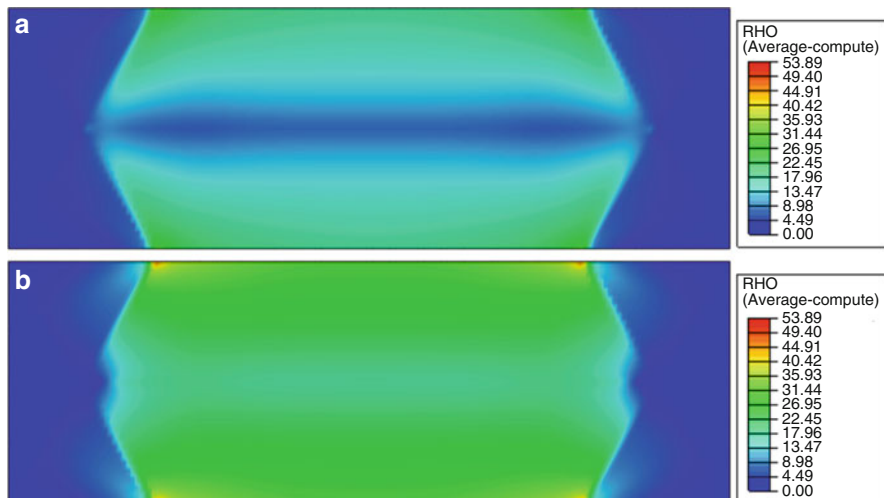


Fig. 14 Dislocation density distributions predicted by the micropolar crystal plasticity simulations at $\Theta = 0.02$ for $\vartheta = 60^\circ$: (a) $H = 8 \mu\text{m}$ (b) $H = 4 \mu\text{m}$

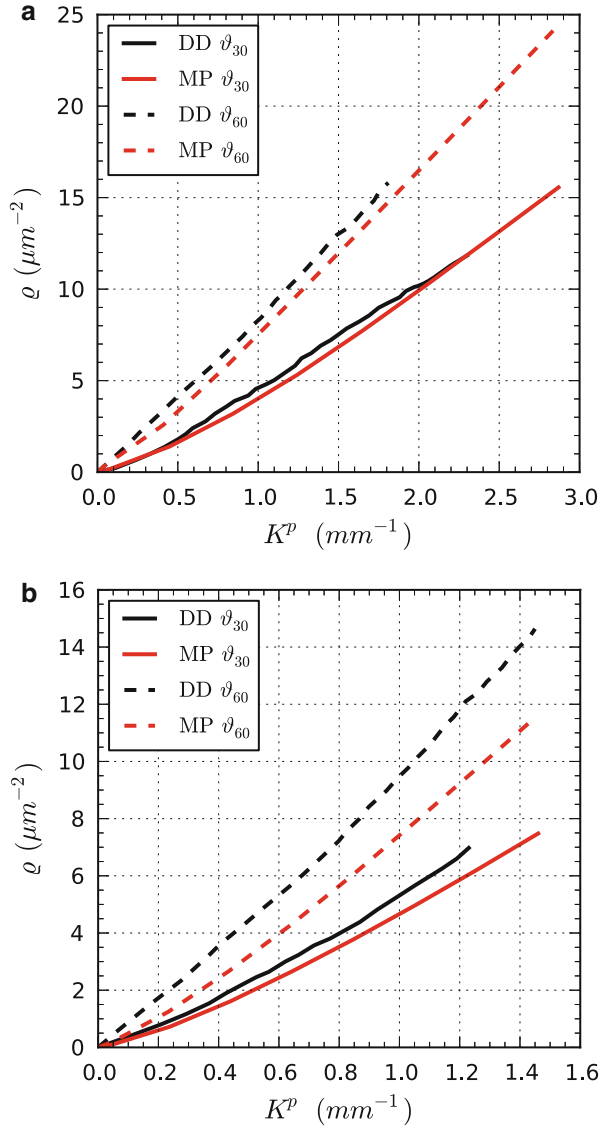
density fields are comparable for both orientations, thus the total (over the entire plastic region) dislocation density is much higher for the ϑ_{60} film.

The total film dislocation density is plotted for both slip orientations and thicknesses versus the average plastic curvature in Fig. 15. In general, the micropolar results compare favorably to the discrete dislocation results. The micropolar model captures the change in slope of the dislocation density-plastic curvature plot with the change in thickness for the ϑ_{30} oriented films, but not for the ϑ_{60} oriented films. An increase in the slope is predicted for the ϑ_{30} films but the micropolar results are nearly identical for the ϑ_{60} films. The DDD results for the $8 \mu\text{m}$ thick film show an increase in slope with increasing average plastic curvature, while the slope for the $4 \mu\text{m}$ remains essentially constant.

Simple Shear of a Metal Matrix Composite

In this application we study the size-dependent hardening of a metal matrix composite subjected to simple shear. The idealized particle reinforced system shown in Fig. 16 was previously analyzed using DDD simulations by Cleveringa et al. (1997, 1999) and Yefimov et al. (2004a) and micropolar crystal plasticity by Mayeur and McDowell (2015). The periodic unit cell consists of an elastic-viscoplastic matrix phase (white) with a single slip system parallel to the x_1 direction, reinforced by elastic particles (gray). The size of the unit cell is $2W \times 2H$ ($W = H\sqrt{3}$) and the particles have dimensions $2W_f \times 2H_f$. Two distinct cases, denoted Material I and Material II, with different particle aspect ratios but the same area fraction, $A_f = (H_f W_f)/(HW) = 0.2$, are studied: the unit cell for Material I is reinforced

Fig. 15 Comparison of the discrete dislocation (DD) and micropolar crystal plasticity (MP) dislocation density evolution as a function of average plastic curvature for the double slip configuration (a) $H = 4 \mu\text{m}$ and (b) $H = 8 \mu\text{m}$. The discrete dislocation results are from Yefimov and van der Giessen (2005a)



by square particles with $W_f = H_f = 0.416H$, while the unit cell for Material II is reinforced by rectangular particles with $H_f = 2W_f = 0.588H$. The two cases are differentiated such that Material I contains an unobstructed vein of matrix material that spans the unit cell, whereas the particles overlap in Material II and block slip. Since the area fraction of the elastic particles is the same for both morphologies, any observed differences in material response are due to the dislocation-particle interactions and not the phase volume fraction.

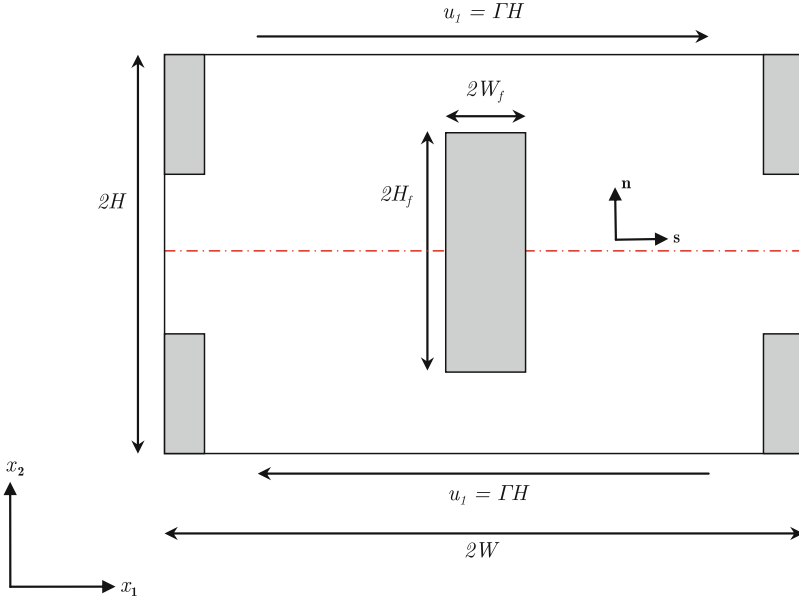


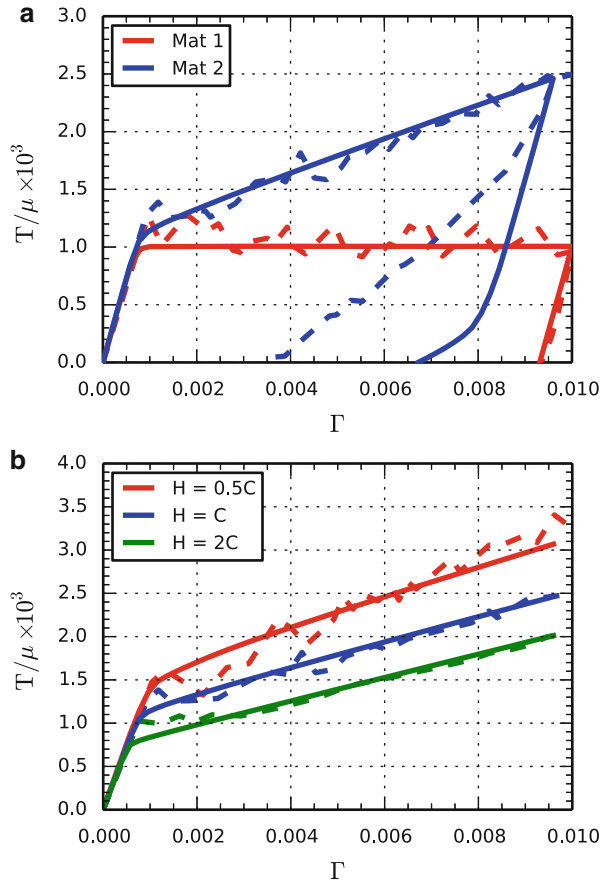
Fig. 16 Schematic of the geometry and slip system configuration for the metal matrix composite initial-boundary value problem

The composite is subjected to simple shear through displacement boundary conditions applied to top and bottom surfaces and these surfaces are assumed to be couple stress traction free. Periodic boundary conditions are enforced on displacements and microrotation at the left and right surfaces. The deformation is imposed at a shearing rate of $\dot{\Gamma} = 10^{-3} \text{s}^{-1}$ up to a unit cell shear strain of $\Gamma = 0.01$, at which point the material is unloaded back to zero strain. The boundary conditions are stated as:

$$\begin{aligned}
 u_1(x_1, \pm H, t) &= \pm \Gamma(t)H, & u_2(x_1, \pm H, t) &= 0 \\
 u_1(-W, x_2, t) &= u_1(W, x_2, t), & u_2(-W, x_2, t) &= u_2(W, x_2, t) \\
 \bar{\phi}(-W, x_2, t) &= \bar{\phi}(W, x_2, t).
 \end{aligned}
 \tag{115}$$

Further, microrotation is assumed to be coupled at the matrix-particle interface. Simulations are performed for cell sizes of $H = \{0.5C, C, 2C\}$ where $C = 4000b$ with $b = 0.25 \text{ nm}$. The FE mesh consists of 106×61 bilinear quadrilateral elements. The classical elastic properties of the particles are $\mu = 192.3 \text{ GPa}$ and $\nu = 0.17$, and the matrix constitutive parameters are listed in Table 1. These parameters are representative of silicon carbide particles embedded in an aluminum matrix. With regard to specifying the nonclassical elastic constants for inclusion, μ_c^I and β^I , we assume that $\mu_c^I = \mu_c^M$ and $\beta^I = \beta^M$ where the superscripts I and M refer to the inclusion and matrix, respectively. As discussed by Cordero et al. (2010), other choices are possible and perhaps should be considered in future work.

Fig. 17 Average stress-strain response for (a) Material I and II for $H = C$ and (b) Material II with variable slip threshold for different particle spacings. Dashed lines are discrete dislocation results (Yefimov 2004a) and solid lines are micropolar results



As shown in Fig. 17a, the unit-cell average stress-strain response predicted by the discrete dislocation model for Material I is nearly elastic-perfectly plastic, whereas Material II displays an approximately linear hardening rate. Material I does not harden because there are no obstacles to dislocation motion. In contrast, mobile dislocations in Material II are obstructed by the particles and form pileups and tilt walls at the matrix-particle interface. The unloading curve suggests the material strengthening in Material II is governed by the development of a strong back stress, as evidenced by the pronounced Bauschinger effect, while unloading is essentially elastic for Material I. The average stress-strain curves predicted by the micropolar model (solid lines in Fig. 17) are in good agreement with the discrete dislocation results during forward loading; however, the Bauschinger effect for Material II is significantly underestimated. Figure 17b shows the stress-strain curves for Material II for the three unit cell sizes. We found that it was necessary to use different r_0 values to obtain good agreement with the discrete dislocation results. Interestingly, the calibrated values, $r_0 = \{13, 21, 30\}$ MPa, show strong correlation with the Hall-

Petch relation $r_0 \propto \Lambda^{-1/2}$, where $\Lambda = (2\sqrt{3} - 0.588)H$ is the mean free path for Material II. The slope of the Hall-Petch relation is $40.51 \text{ MPa } \sqrt{\mu\text{m}}$. Note that the Hall-Petch relation was not assumed a priori, but rather is consistent with the result of parameter estimation.

The cumulative plastic slip distributions for both cases with $H = C$ and $\Gamma = 0.006$ as predicted by the micropolar model are shown in Fig. 18. The slip field morphology for Material I is characterized by intensely localized plasticity in the unreinforced veins of matrix material, whereas the slip morphology for Material II is characterized by highly localized bands that form along the top and bottom faces of the particles, but do not extend across the full width of the unit cell due to the particle overlap. The cumulative slip distributions predicted by the micropolar model are consistent with the discrete dislocation simulations and are noticeably different than the predictions of local and low-order gradient theories, which show much higher levels of slip accumulation along the vertical matrix-particle interfaces.

The total dislocation density fields are plotted for all three cases of Material II in Fig. 19 at $\Gamma = 0.006$. There is a significant dislocation density accumulation

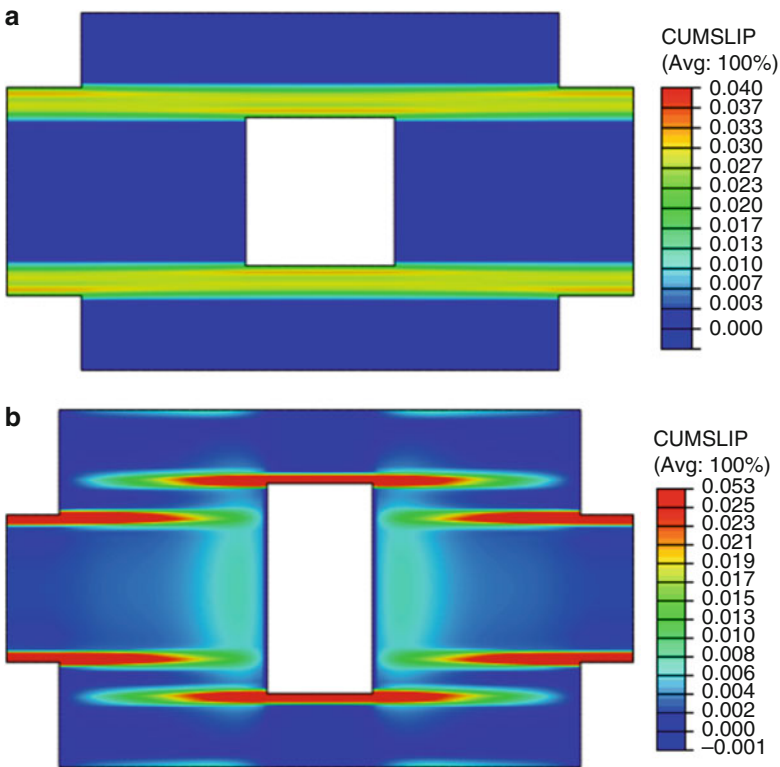


Fig. 18 Contours of cumulative slip at $\Gamma = 0.006$ for (a) Material I (b) Material II ($H = C$)

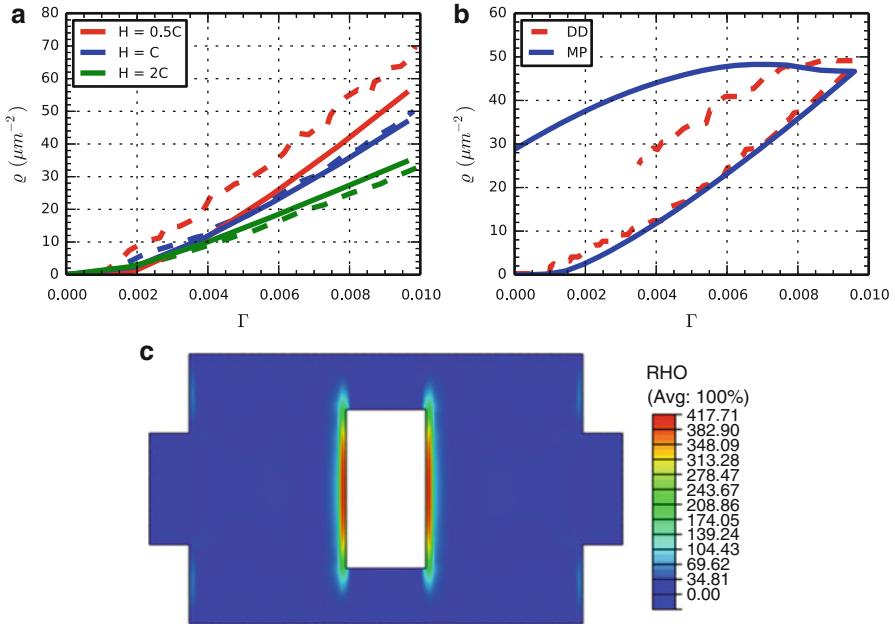


Fig. 19 Evolution of total matrix dislocation density versus applied strain as predicted by the micropolar model (solid lines) and discrete dislocation dynamics (dashed lines) (Yefimov et al. 2004a) (a) for various unit cell sizes during loading and (b) unloading for $H = C$. Contour of total dislocation density for Material II at $\Gamma = 0.006$ for the micropolar model $H = 0.5C$

along the vertical faces of the matrix-particle interface and sparse dislocation density distributed throughout the matrix. These GNDs are generated to accommodate the rotational gradients that develop at the interface (Ashby 1970). The total matrix dislocation density is plotted during loading as a function of applied shear strain for both the micropolar and discrete dislocation models for the three unit cell sizes in Fig. 19a and during loading/unloading for $H = C$ in Fig. 19b. The micropolar results are in good agreement with the discrete dislocation simulations for the two largest unit cells during loading, but there is a modest departure in the model predictions for the $H = 0.5C$ case. Interestingly, the discrete dislocation model predicts much higher rate of dislocation recovery upon unloading as compared to the micropolar model.

Overall, the predictions of the micropolar simulations are in reasonably good agreement with the discrete dislocation simulation results. The most significant discrepancy in the simulated material response is the underprediction of the Bauschinger effect and the rate of dislocation recovery upon unloading. We believe these discrepancies are largely related to the higher-order boundary conditions enforced at the matrix-particle interface. It is assumed that the lattice rotations at the matrix-particle interface are equal in these simulations, i.e., the finite element nodes along the matrix-particle interface are shared between the two materials. This

represents a different boundary condition than what is enforced at the matrix-particle interface in the discrete dislocation simulation. The interface boundary condition enforced in the DDD models is one of equal displacements and zero slip at the vertical matrix-particle faces. Thus, the micropolar model enforces an additional constraint at the matrix-particle interface, which may overconstrain the material response upon unloading. It may be possible to achieve an improved unloading response by either modifying boundary conditions at the matrix-particle interface or the nonclassical elastic constants of the inclusion, μ_c^I and β^I , which will alter the way higher-order tractions are transmitted between the two phases.

Conclusions

Micropolar crystal plasticity is a specialized subset of the more general micromorphic crystal plasticity theory, discussed in a subsequent chapter, that accounts for size effects due to gradients of lattice rotation. The connection between the micropolar lattice torsion-curvature and Nye's GND tensor was established and related to concepts from slip gradient-based crystal plasticity. The advantage of the micropolar theory in comparison to the micromorphic and slip gradient-based theories is the reduced complexity in that it requires only three additional continuum degrees of freedom for the fully three-dimensional case.

A full treatment of the finite deformation kinematics and thermodynamic-based constitutive equations have been developed and placed in context as an extension of concepts of local crystal plasticity theory. A model employing linearized kinematics is then presented with an explicit set of constitutive equations that were used in finite element simulations of initial-boundary value problems previously solved using discrete dislocation dynamics (DDD). The simulation results demonstrate the ability of the micropolar theory to capture many of the salient features exhibited by the DDD simulations for a wide range of boundary value problems, including both the size-dependence of the stress-strain response and the evolution of the dislocation density.

An extended comparison to Gurtin-type slip gradient-based theories of higher-order single crystal plasticity was presented. The analysis highlights many striking theoretical similarities, which suggests that they will also share many of the same predictive capabilities. A few subtle, but key differences, with respect to the construction of dissipative constitutive equations are also discussed. For example, it is possible in the slip gradient-based theories to isolate gradient energetic and dissipative length scale effects, whereas they are coupled in the micropolar theory, i.e., one cannot have gradient dissipative effects in the absence of energetic gradient effects since there will be no driving force for plastic torsion-curvature evolution. However, it is noted that there is no principal restriction preventing one from casting the micropolar theory in terms of energetic-dissipative decomposition of the couple stress tensor rather than of the lattice torsion-curvature, which would then accommodate a true separation of energetic and dissipative gradient effects.

There is still much exciting work to be done in further developing the micropolar theory. Two areas of particular interest are the development of proper intermediate higher-order boundary conditions between fully constrained (micro-hard) and traction free (micro-free) and the proper description of interface boundary conditions and/or constitutive equations for grain and/or phase boundaries.

Acknowledgments JRM acknowledges the support of Los Alamos National Laboratory, operated by Los Alamos National Security LLC under DOE Contract DE-AC52-06NA25936. This work benefited from the support of the Laboratory Directed Research and Development Early Career award 20150696ECR. DLM would like to acknowledge the support of the Carter N. Paden, Jr. Distinguished Chair in Metals Processing.

References

- Abaqus/Standard Version 6.7.1. Dassault Systèmes Simulia Corp. (2007)
- E. Aifantis, On the microstructural origin of certain inelastic models. *J. Eng. Mater. Technol.* **106**, 326–330 (1984)
- A. Arsenlis, D.M. Parks, Crystallographic aspects of geometrically-necessary and statistically-stored dislocation density. *Acta Mater.* **47**, 1597–1611 (1999)
- M.F. Ashby, Deformation of plastically non-homogeneous materials. *Philos. Mag.* **21**(170), 399–424 (1970)
- E. Bittencourt, A. Needleman, M.E. Gurtin, E. van der Giessen, A comparison of nonlocal continuum and discrete dislocation plasticity predictions. *J. Mech. Phys. Solids* **51**, 281–310 (2003)
- P. Cermelli, M. Gurtin, On the characterization of geometrically necessary dislocations in finite plasticity. *J. Mech. Phys. Solids* **49**, 1539–1568 (2001)
- J. Clayton, D. Bamman, D. McDowell, A geometric framework for the kinematics of crystals with defects. *Philos. Mag.* **85**, 3983–4010 (2005)
- H.H.M. Cleveringa, E. van der Giessen, A. Needleman, Comparison of discrete dislocation and continuum plasticity predictions for a composite material. *Acta Mater.* **45**, 3163–3179 (1997)
- H.H.M. Cleveringa, E. van der Giessen, A. Needleman, A discrete dislocation analysis of residual stresses in a composite material. *Philos. Mag. A Phys. Condens. Matter Struct. Defects Mech Prop* **79**, 893–920 (1999)
- B.D. Coleman, M.E. Gurtin, Thermodynamics with internal state variables. *J. Chem. Phys.* **47**(2), 597–613 (1967)
- N.M. Cordero, A. Gaubert, S. Forest, E.P. Busso, F. Gallerneau, S. Kruch, Size effects in generalised continuum crystal plasticity for two-phase laminates. *J. Mech. Phys. Solids* **58**, 1963–1994 (2010)
- A. Eringen, *Microcontinuum Field Theories I: Foundations and Solids* (Springer, New York, 1999)
- A.C. Eringen, W.D. Claus Jr., A micromorphic approach to dislocation theory and its relation to several existing theories. Technical Report TR-6, Princeton University Department of Aerospace and Mechanical Sciences (1969)
- A.C. Eringen, E.S. Suhubi, Nonlinear theory of simple micro-elastic solids-I. *Int. J. Eng. Sci.* **2**, 189–203 (1964)
- L.P. Evers, W.A.M. Brekelmans, M.G.D. Geers, Non-local crystal plasticity model with intrinsic SSD and GND effects. *J. Mech. Phys. Solids* **52**, 2379–2401 (2004)
- S. Forest, Some links between Cosserat, strain gradient crystal plasticity and the statistical theory of dislocations. *Philos. Mag.* **88**, 3549–3563 (2008)
- S. Forest, Generalized continuum modelling of crystal plasticity, in *Generalized Continua and Dislocation Theory* (Springer, Vienna, 2012), pp. 181–287

- S. Forest, R. Sievert, Elastoviscoplastic constitutive frameworks for generalized continua. *Acta Mech.* **160**, 71–111 (2003)
- S. Forest, G. Cailletaud, R.W. Sievert, A Cosserat theory for elastoviscoplastic single crystals at finite deformation. *Arch. Mech.* **49**, 705–736 (1997)
- S. Forest, F. Barbe, G. Cailletaud, Cosserat modelling of size effects in the mechanical behaviour of polycrystals and multi-phase materials. *Int. J. Solids Struct.* **37**(46), 7105–7126 (2000)
- P. Germain, The method of virtual power in continuum mechanics. Part 2: microstructure. *SIAM J. Appl. Math.* **25**, 556–575 (1973)
- W. Günther, Zur Statik und Kinematik des Cosseratschen Kontinuums. *Abh. Braunsch. Wiss. Ges.* **10**, 195–213 (1958)
- M.E. Gurtin, A gradient theory of single-crystal viscoplasticity that accounts for geometrically necessary dislocations. *J. Mech. Phys. Solids* **50**, 5–32 (2002)
- M.E. Gurtin, L. Anand, S.P. Lele, Gradient single-crystal plasticity with free energy dependent on dislocation densities. *J. Mech. Phys. Solids* **55**, 1853–1878 (2007)
- E. Kröner, On the physical reality of torque stresses in continuum mechanics. *Int. J. Eng. Sci.* **1**, 261–278 (1963)
- E. Kröner, Initial studies of a plasticity theory based upon statistical mechanics, in *Inelastic Behaviour of Solids*, ed. by M. Kanninen, W. Adler, A. Rosenfield, R. Jaffee (McGraw-Hill, New York, 1969), pp. 137–147
- M. Kuroda, V. Tvergaard, A finite deformation theory of higher-order gradient crystal plasticity. *J. Mech. Phys. Solids* **56**(8), 2573–2584 (2008)
- S. Limkumnerd, E. van der Giessen, Study of size effects in thin films by means of a crystal plasticity theory based on DiFT. *J. Mech. Phys. Solids* **56**, 3304–3314 (2008)
- J.R. Mayeur, Generalized continuum modeling of scale-dependent crystalline plasticity. Ph.D. thesis, Georgia Institute of Technology, 2010
- J.R. Mayeur, D.L. McDowell, Bending of single crystal thin films modeled with micropolar crystal plasticity. *Int. J. Eng. Sci.* **49**, 1357–1366 (2011)
- J.R. Mayeur, D.L. McDowell, An evaluation of higher-order single crystal strength models for constrained thin films subjected to simple shear. *J. Mech. Phys. Solids* **61**, 1935–1954 (2013)
- J.R. Mayeur, D.L. McDowell, A comparison of Gurtin type and micropolar theories of generalized single crystal plasticity. *Int. J. Plast.* **57**, 29–51 (2014)
- J.R. Mayeur, D.L. McDowell, Micropolar crystal plasticity simulation of particle strengthening. *Model. Simul. Mater. Sci. Eng.* **23**, 065007 (2015)
- J.R. Mayeur, D.L. McDowell, D.J. Bammann, Dislocation-based micropolar single crystal plasticity: comparison of multi- and single criterion theories. *J. Mech. Phys. Solids* **59**, 398–422 (2011)
- H. Mecking, U.F. Kocks, Kinetics of flow and strain-hardening. *Acta Metall.* **29**, 1865–1875 (1981)
- J.F. Nye, Some geometrical relations in dislocated crystals. *Acta Metall.* **1**, 153–162 (1953)
- B.D. Reddy, The role of dissipation and defect energy in variational formulations of problems in strain-gradient plasticity. Part 2: single-crystal plasticity. *Contin. Mech. Thermodyn.* **23**, 551–572 (2011)
- H. Schäfer, Eine Feldtheorie der Versetzungen im Cosserat-Kontinuum. *Z. Angew. Math. Phys.* **20**, 891–899 (1969)
- J.Y. Shu, N.A. Fleck, E. van der Giessen, A. Needleman, Boundary layers in constrained plastic flow: comparison of nonlocal and discrete dislocation plasticity. *J. Mech. Phys. Solids* **49**, 1361–1395 (2001)
- R. Sievert, S. Forest, R. Trostel, Finite deformation Cosserat-type modelling of dissipative solids and its application to crystal plasticity. *J. Phys.* **IV 8**, 357–364 (1998)
- C. Teodosiu, F. Sidoroff, A theory of finite elastoviscoplasticity of single crystals. *Int. J. Eng. Sci.* **14**, 165–176 (1976)
- S. Yefimov, E. van der Giessen, Multiple slip in a strain-gradient plasticity model motivated by a statistical-mechanics description of dislocations. *Int. J. Solids Struct.* **42**, 3375–3394 (2005a)
- S. Yefimov, E. van der Giessen, Size effects in single crystal thin films: nonlocal crystal plasticity simulations. *Eur. J. Mech. A Solids* **24**, 183–193 (2005b)

-
- S. Yefimov, I. Groma, E. van der Giessen, A comparison of a statistical-mechanics based plasticity model with discrete dislocation plasticity calculations. *J. Mech. Phys. Solids* **52**, 279–300 (2004a)
- S. Yefimov, E. van der Giessen, I. Groma, Bending of a single crystal: discrete dislocation and nonlocal crystal plasticity simulations. *Model. Simul. Mater. Sci. Eng.* **12**, 1069 (2004b)



Samuel Forest, J. R. Mayeur, and D. L. McDowell

Contents

Introduction	644
The Microcurl Model at Finite Deformation	645
Model Formulation	645
Geometrically Linearized Model	650
Comparison Between Micropolar and Micromorphic Crystal Plasticity	651
Size Effects in a Two-Phase Single-Crystal Laminate	653
Strain Gradient Plasticity as a Limit Case	659
Free Energy Potentials for Micromorphic Crystal Plasticity	662
Formulation of Two Free Energy Potentials	663
Application to the Shearing of the Periodic Laminate	665
Cyclic Behavior of the Laminate	670
Grain Size Effects in Polycrystals	674
Boundary Value Problem for Polycrystals	675
Overall Cyclic Response of a Polycrystalline Aggregate	676
Grain Size Effects in Idealized Aluminum Polycrystals	679
Conclusions	682
References	683

S. Forest (✉)
Centre des Matériaux, Mines ParisTech CNRS, PSL Research University, CNRS UMR 7633,
Evry Cedex, Paris, France
e-mail: samuel.forest@mines-paristech.fr

J. R. Mayeur
Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM, USA
e-mail: jason.mayeur@gmail.com

D. L. McDowell
Woodruff School of Mechanical Engineering, School of Materials Science and Engineering,
Georgia Institute of Technology, Atlanta, GA, USA
e-mail: david.mcdowell@me.gatech.edu

Abstract

The micromorphic approach to crystal plasticity represents an extension of the micropolar (Cosserat) framework, which is presented in a separate chapter. Cosserat crystal plasticity is contained as a special constrained case in the same way as the Cosserat theory is a special restricted case of Eringen's micromorphic model, as explained also in a separate chapter. The micromorphic theory is presented along the lines of Aslan et al. (Int J Eng Sci 49:1311–1325, 2011) and Forest et al. (Micromorphic approach to crystal plasticity and phase transformation. In: Schroeder J, Hackl K (eds) Plasticity and beyond. CISM international centre for mechanical sciences, courses and lectures, vol 550, Springer, pp 131–198, 2014) and compared to the micropolar model in some applications. These extensions of conventional crystal plasticity aim at incorporating the dislocation density tensor introduced by Kröner (Initial studies of a plasticity theory based upon statistical mechanics. In: Kanninen M, Adler W, Rosenfield A, Jaffee R (eds) Inelastic behaviour of solids. McGraw-Hill, pp 137–147, 1969). and Cermelli and Gurtin (J Mech Phys Solids 49:1539–1568, 2001) into the constitutive framework. The concept of dislocation density tensor is equivalent to that of the so-called geometrically necessary dislocations (GND) introduced by Ashby (The deformation of plastically non-homogeneous alloys. In: Kelly A, Nicholson R (eds) Strengthening methods in crystals. Applied Science Publishers, London, pp 137–192, 1971). The applications presented in this chapter deal with pile-up formation in laminate microstructures and strain localization phenomena in polycrystals.

Keywords

Micromorphic medium · Crystal plasticity · Dislocation density tensor · Geometrically necessary dislocations · Strain gradient plasticity · Size effect

Introduction

The micromorphic approach to crystal plasticity represents an extension of the micropolar (Cosserat) framework which is presented in a separate chapter. Cosserat crystal plasticity is contained as a special constrained case in the same way as the Cosserat theory is a special restricted case of Eringen's micromorphic model, as explained also in a separate chapter. The micromorphic theory is presented along the lines of Aslan et al. (2011) and Forest et al. (2014) and compared to the micropolar model in some applications. These extensions of conventional crystal plasticity aim at incorporating the dislocation density tensor introduced by Kröner (1969) and Cermelli and Gurtin (2001) into the constitutive framework. The concept of dislocation density tensor is equivalent to that of so-called geometrically necessary dislocations (GND) introduced by Ashby (1971).

The links between the micromorphic continuum and the plasticity of crystalline materials have been recognized very early by Claus and Eringen (1969) and Eringen

and Claus (1970). Lattice directions in a single crystal can be regarded as directors that rotate and deform as they do in a micromorphic continuum. The fact that lattice directions can be rotated and stretched in a different way than material lines connecting individual atoms, especially in the presence of static or moving dislocations, illustrates the independence between directors and material lines in a micromorphic continuum, even though their deformation can be related at the constitutive level.

The objective of the present chapter is to formulate a finite deformation micromorphic extension of conventional crystal plasticity to account for GND effects in single crystals. It also provides analytical predictions of size effects on the yield strength and kinematic hardening of laminate microstructures made of an elastic layer and an elastic–plastic single-crystal layer undergoing single slip. The theory is called the *microcurl* model because the evaluation of the curl of the microdeformation, instead of its full gradient, is sufficient to account for the effect of the dislocation density tensor.

The models proposed in this section for single crystals fall in the class of anisotropic elastoviscoplastic micromorphic media for which constitutive frameworks at finite deformations have been proposed in Forest and Sievert (2003), Lee and Chen (2003), Grammenoudis and Tsakmakis (2009), Sansour et al. (2010), and Regueiro 2010; see the corresponding chapter in this handbook. In fact, the micromorphic approach can be applied not only to the total deformation by introducing the microdeformation field but can also be restricted to plastic deformation, for specific application to size effects in plasticity, or to damage variables for application to regularized simulation of crack propagation, as proposed in Forest (2009, 2016) and Hirschberger and Steinmann (2009).

The outline of this chapter is as follows. The crystal plasticity model formulated within Eringen’s micromorphic framework is presented at finite deformation in section “[The Microcurl Model at Finite Deformation](#),” together with its linearization. Size effects predicted by the model are illustrated in section “[Size Effects in a Two-Phase Single-Crystal Laminate](#).” Some constitutive laws involving the dislocation density tensor are discussed in section “[Free Energy Potentials for Micromorphic Crystal Plasticity](#)” with an application to cyclic plasticity in single crystals. Finally, the model is used to predict the response of polycrystalline metals and alloys in section “[Grain Size Effects in Polycrystals](#).”

The Microcurl Model at Finite Deformation

Model Formulation

Balance Equations

The degrees of freedom of the proposed theory are the displacement vector \mathbf{u} and the microdeformation variable $\widehat{\chi}^p$, a generally nonsymmetric second-rank tensor. The field $\widehat{\chi}^p(\mathbf{X})$ is generally not compatible, meaning that it does not derive from a vector field. The exponent p indicates, in advance, that this variable will eventually

be constitutively related to plastic deformation occurring at the material point. In particular, the microdeformation $\widehat{\chi}^p$ is treated as an invariant quantity with respect to rigid body motion. The polar decomposition of the microdeformation contains the polar rotation $\overline{\mathbf{R}}$ used in the micropolar crystal plasticity theory and a symmetric microstretch tensor. As a result, when this microstretch tensor is close to the identity tensor, the micromorphic model reduces to the micropolar one.

A first gradient theory is considered with respect to the degrees of freedom. However, the influence of the microdeformation gradient is limited to its curl part because of the intended relation to the dislocation density tensor associated with the curl of plastic distortion. The following sets of degrees of freedom and of their gradients are therefore defined:

$$DOF = \{ \mathbf{u}, \widehat{\chi}^p \}, GRAD = \{ \mathbf{F} := \mathbf{1} + \mathbf{u} \otimes \nabla_0, \mathbf{K} := \text{Curl } \widehat{\chi}^p \} \quad (1)$$

The following definition of the curl operator is adopted:

$$\text{Curl } \widehat{\chi}^p := \frac{\partial \widehat{\chi}^p}{\partial X_k} \times \mathbf{e}_k, K_{ij} := \epsilon_{jkl} \frac{\partial \widehat{\chi}^p_{ik}}{\partial X_l} \quad (2)$$

where ϵ_{ijk} is the permutation tensor.

The method of virtual power is used to derive the balance and boundary conditions, following Germain (1973). For that purpose, the power density of internal forces is defined as a linear form with respect to the velocity fields and their Eulerian gradients:

$$p^{(i)} = \boldsymbol{\sigma} : (\dot{\mathbf{u}} \otimes \nabla) + \mathbf{s} : \dot{\widehat{\chi}}^p + \mathbf{M} : \text{curl } \dot{\widehat{\chi}}^p, \forall \mathbf{x} \in V \quad (3)$$

Here, the conjugate quantities are the Cauchy stress tensor $\boldsymbol{\sigma}$, which is symmetric for objectivity reasons; the microstress tensor, \mathbf{s} ; and the generalized couple-stress tensor \mathbf{M} . The curl of the microdeformation rate is defined as:

$$\text{curl } \dot{\widehat{\chi}}^p := \epsilon_{jkl} \frac{\partial \dot{\widehat{\chi}}^p_{ik}}{\partial X_l} \mathbf{e}_i \otimes \mathbf{e}_j = \dot{\mathbf{K}} \mathbf{F}^{-1} \quad (4)$$

The form of the power density of internal forces dictates the form of the power density of contact forces:

$$p^{(c)} = \mathbf{t} \cdot \dot{\mathbf{u}} + \mathbf{m} : \dot{\widehat{\chi}}^p, \forall \mathbf{x} \in \partial V \quad (5)$$

where \mathbf{t} is the usual simple traction vector and \mathbf{m} is the double-traction tensor. The principle of virtual power is stated in the static case and in the absence of volume forces for the sake of brevity:

$$-\int_D p^{(i)} dV + \int_{\partial D} p^{(c)} dS = 0 \quad (6)$$

for all virtual fields $\dot{\mathbf{u}}, \hat{\chi}^p$ and any subdomain $D \subset V$. By application of the Gauss divergence theorem, assuming sufficient regularity of the fields, this statement expands into:

$$\int_V \frac{\partial \sigma_{ij}}{\partial x_j} \dot{u}_i dV + \int_V \left(\epsilon_{kjl} \frac{\partial M_{ik}}{\partial x_l} - s_{ij} \right) \hat{\chi}_{ij}^p dV + \int_{\partial V} (t_i - \sigma_{ij} n_j) \dot{u}_i dS + \int_{\partial V} (m_{ik} - \epsilon_{jkl} M_{ij} n_l) \hat{\chi}_{ik}^p dS = 0, \quad \forall \dot{u}_i, \forall \hat{\chi}_{ij}^p$$

which leads to the two-field equations of balance of momentum and generalized balance of moment of momentum:

$$\operatorname{div} \boldsymbol{\sigma} = 0, \operatorname{curl} \mathbf{M} + \mathbf{s} = 0, \quad \forall \mathbf{x} \in V \quad (7)$$

and two boundary conditions:

$$\mathbf{t} = \boldsymbol{\sigma} \cdot \mathbf{n}, \mathbf{m} = \mathbf{M} \cdot \boldsymbol{\epsilon} \cdot \mathbf{n}, \quad \forall \mathbf{x} \in \partial V \quad (8)$$

the index representation of the latter relation being $m_{ij} = M_{ik} \epsilon_{kjl} n_l$. These balance equations can be compared to the corresponding ones in the chapter dedicated to the micropolar theory.

Constitutive Equations

The deformation gradient is decomposed into elastic and plastic parts in the form,

$$\mathbf{F} = \mathbf{F}^e \mathbf{F}^p \quad (9)$$

The isoclinic intermediate configuration is defined in a unique way by keeping the crystal orientation unchanged from the initial to the intermediate configuration following Mandel (1973). The plastic distortion \mathbf{F}^p is invariant with respect to rigid body motions that are carried by \mathbf{F}^e . The current mass density is ρ , whereas the mass density of the material element in the intermediate configuration is $\tilde{\rho}$, such that $\tilde{\rho}/\rho = J_e := \det(\mathbf{F}^e)$. The elastic strain is defined as:

$$\tilde{\mathbf{E}}^e := \frac{1}{2} (\mathbf{F}^{eT} \mathbf{F}^e - \mathbf{1}) \quad (10)$$

The microdeformation is linked to the plastic deformation via the introduction of a relative deformation measure, defined as:

$$\mathbf{e}^p := \mathbf{F}^{p-1} \hat{\chi}^p - \mathbf{1} \quad (11)$$

This tensor \mathbf{e}^p measures the departure of the microdeformation from the plastic deformation. The state variables are assumed to be the elastic strain, the relative deformation, the curl of microdeformation, and some internal variables, α :

$$STATE := \left\{ \tilde{\mathbf{E}}^e, \mathbf{e}^p, \mathbf{K}, \alpha \right\} \quad (12)$$

The specific Helmholtz free energy density, ψ , is assumed to be a function of this set of state variables. In particular, in this simple version of the model, the curl of microdeformation is assumed to contribute entirely to the stored energy. In more sophisticated models, as proposed in Forest and Sievert (2003, 2006), Forest (2009), and Gurtin and Anand (2009), the relative deformation, the microdeformation, and its gradient can be split into elastic and plastic parts.

When the internal constraint $\mathbf{e}^p \equiv 0$ is enforced, the plastic microdeformation coincides with the plastic deformation so that the curl of the plastic microdeformation is directly related to the dislocation density tensor previously defined by:

$$\mathbf{K} := \text{Curl } \widehat{\boldsymbol{\chi}}^p \equiv \text{Curl } \mathbf{F}^p = J \mathbf{A} \mathbf{F}^{-T} \quad (13)$$

where \mathbf{A} is the dislocation density tensor defined as the curl of the inverse elastic deformation.

The micromorphic model then reduces to strain gradient plasticity according to Gurtin (2002).

The dissipation rate density is the difference:

$$\Delta^{intr} := p^{(i)} - \rho \dot{\psi} \geq 0 \quad (14)$$

which must be positive according to the second principle of thermodynamics. When the previous strain measures are introduced, the power density of internal forces takes the following form:

$$\begin{aligned} p^{(i)} &= \boldsymbol{\sigma} : \dot{\mathbf{F}}^e \mathbf{F}^{e-1} + \boldsymbol{\sigma} : \left(\mathbf{F}^e \dot{\mathbf{F}}^p \mathbf{F}^{p-1} \mathbf{F}^{e-1} \right) + \mathbf{s} : \left(\mathbf{F}^p \dot{\mathbf{e}}^p + \dot{\mathbf{F}}^p \mathbf{e}^p \right) + \mathbf{M} : \dot{\mathbf{K}} \mathbf{F}^{-1} \\ &= \frac{\rho}{\tilde{\rho}} \boldsymbol{\Pi}^e : \dot{\tilde{\mathbf{E}}}^e + \frac{\rho}{\tilde{\rho}} \boldsymbol{\Pi}^M : \dot{\mathbf{F}}^p \mathbf{F}^{p-1} + \mathbf{s} : \left(\mathbf{F}^p \dot{\mathbf{e}}^p + \dot{\mathbf{F}}^p \mathbf{e}^p \right) + \mathbf{M} : \dot{\mathbf{K}} \mathbf{F}^{-1} \end{aligned} \quad (15)$$

where $\boldsymbol{\Pi}^e$ is the second Piola–Kirchhoff stress tensor with respect to the intermediate configuration and $\boldsymbol{\Pi}^M$ is the Mandel stress tensor:

$$\boldsymbol{\Pi}^e := J_e \mathbf{F}^{e-1} \boldsymbol{\sigma} \mathbf{F}^{e-T}, \boldsymbol{\Pi}^M := J_e \mathbf{F}^{eT} \boldsymbol{\sigma} \mathbf{F}^{e-T} = \mathbf{F}^{eT} \mathbf{F}^e \boldsymbol{\Pi}^e \quad (16)$$

On the other hand,

$$\rho \dot{\psi} = \rho \frac{\partial \psi}{\partial \tilde{\mathbf{E}}^e} : \dot{\tilde{\mathbf{E}}}^e + \rho \frac{\partial \psi}{\partial \mathbf{e}^p} : \dot{\mathbf{e}}^p + \rho \frac{\partial \psi}{\partial \mathbf{K}} : \dot{\mathbf{K}} + \rho \frac{\partial \psi}{\partial \alpha} \dot{\alpha} \quad (17)$$

We compute:

$$\begin{aligned}
 J_e D &= \left(\boldsymbol{\Pi}^e - \tilde{\rho} \frac{\partial \psi}{\partial \mathbf{E}^e} \right) : \dot{\mathbf{E}}^e + \left(J_e \mathbf{F}^{pT} \mathbf{s} - \tilde{\rho} \frac{\partial \psi}{\partial \mathbf{e}^p} \right) : \dot{\mathbf{e}}^p \\
 &+ \left(J_e \mathbf{M} \mathbf{F}^{-T} - \tilde{\rho} \frac{\partial \psi}{\partial \mathbf{K}} \right) : \dot{\mathbf{K}} \\
 &+ \left(\boldsymbol{\Pi}^M + J_e \mathbf{s} \widehat{\boldsymbol{\chi}}^{pT} \right) : \dot{\mathbf{F}}^p \mathbf{F}^{p-1} - \tilde{\rho} \frac{\partial \psi}{\partial \alpha} \dot{\alpha} \geq 0
 \end{aligned} \tag{18}$$

Assuming that the processes associated with $\dot{\mathbf{E}}^e$, $\dot{\mathbf{e}}^p$ and $\dot{\mathbf{K}}$ are nondissipative, the state laws are obtained:

$$\boldsymbol{\Pi}^e = \tilde{\rho} \frac{\partial \psi}{\partial \mathbf{E}^e}, \quad \mathbf{s} = J_e^{-1} \mathbf{F}^{p-T} \tilde{\rho} \frac{\partial \psi}{\partial \mathbf{e}^p}, \quad \mathbf{M} = J_e^{-1} \tilde{\rho} \frac{\partial \psi}{\partial \mathbf{K}} \mathbf{F}^T \tag{19}$$

The residual dissipation rate is:

$$J_e D = \left(\boldsymbol{\Pi}^M + J_e \mathbf{s} \widehat{\boldsymbol{\chi}}^{pT} \right) : \dot{\mathbf{F}}^p \mathbf{F}^{p-1} - R \dot{\alpha} \geq 0, \quad \text{with } R := \tilde{\rho} \frac{\partial \psi}{\partial \alpha} \tag{20}$$

At this stage, a dissipation potential that depends on stress measures, $\Omega(S, R)$, is introduced in order to formulate the evolution equations for plastic flow and internal variables:

$$\dot{\mathbf{F}}^p \mathbf{F}^{p-1} = \frac{\partial \Omega}{\partial \mathcal{S}}, \quad \text{with } \mathcal{S} := \boldsymbol{\Pi}^M + J_e \mathbf{s} \widehat{\boldsymbol{\chi}}^{pT} \tag{21}$$

$$\dot{\alpha} = -\frac{\partial \Omega}{\partial R} \tag{22}$$

where R is the thermodynamic force associated with the internal variable α and S is the effective stress conjugate to plastic strain rate, the driving force for plastic flow.

In the case of crystal plasticity, a generalized Schmid law is adopted for each slip system s in the form:

$$f^s(\mathcal{S}, \tau_c^s) = |\mathcal{S} : \mathbf{P}^s| - \tau_c^s \geq 0, \quad \text{with } \mathbf{P}^s = \mathbf{l}^s \otimes \mathbf{n}^s \tag{23}$$

for activation of slip system s with slip direction, \mathbf{l}^s , and normal to the slip plane, \mathbf{n}^s . We call \mathbf{P}^s the orientation tensor. The critical resolved shear stress is τ_c^s which may be a function of R in the presence of isotropic hardening. The kinematics of plastic slip follows from the choice of a dissipation potential, $\Omega(f^s)$, that depends on the stress variables through the yield function itself, f^s :

$$\dot{\mathbf{F}}^p \mathbf{F}^{p-1} = \sum_{s=1}^N \frac{\partial \Omega}{\partial f^s} \frac{\partial f^s}{\partial \mathcal{S}} = \sum_{s=1}^N \dot{\gamma}^s \mathbf{P}^s, \quad \text{with } \dot{\gamma}^s = \frac{\partial \Omega}{\partial f^s} \text{sign}(\mathcal{S} : \mathbf{P}^s) \tag{24}$$

A possible viscoplastic potential is then:

$$\Omega(f^s) = \frac{K}{n+1} \left\langle \frac{f^s}{K} \right\rangle^{n+1} \quad (25)$$

where K and n are viscosity parameters associated with viscoplastic slip, and the brackets stand for $\langle \cdot \rangle = \text{Max}(0, \cdot)$. The generalized resolved shear stress can be decomposed into two contributions:

$$\mathcal{S} : \mathbf{P}^s = \tau^s - x^s, \text{ with } \tau^s = \boldsymbol{\Pi}^M : \mathbf{P}^s \text{ and } x^s = -\mathbf{s} \widehat{\boldsymbol{\chi}}^{pT} : \mathbf{P}^s \quad (26)$$

The usual resolved shear stress is τ^s , whereas x^s can be interpreted as an internal stress or back stress leading to kinematic hardening. The fact that the introduction of the effect of the dislocation density tensor or, more generally, of gradient of plastic strain tensor leads to the existence of internal stresses induced by higher-order stresses has already been noticed by Steinmann (1996); see also Forest (2008). The back stress component is induced by the microstress \mathbf{s} or, equivalently, by the curl of the generalized couple-stress tensor, \mathbf{M} , via the balance Eq. (7).

Geometrically Linearized Model

When deformations and rotations remain sufficiently small, the previous equations can be linearized as follows:

$$\mathbf{F} = \mathbf{1} + \mathbf{H} \simeq \mathbf{1} + \mathbf{H}^e + \mathbf{H}^p, \mathbf{H}^e = \boldsymbol{\varepsilon}^e + \boldsymbol{\omega}^e, \mathbf{H}^p = \boldsymbol{\varepsilon}^p + \boldsymbol{\omega}^p \quad (27)$$

where $\boldsymbol{\varepsilon}^e$ and $\boldsymbol{\omega}^e$ (resp. $\boldsymbol{\varepsilon}^p$, $\boldsymbol{\omega}^p$) are the symmetric and skew-symmetric parts of $\mathbf{F}^e - \mathbf{1}$ (resp. $\mathbf{F}^p - \mathbf{1}$). When microdeformation is small, the relative deformation is linearized as:

$$\mathbf{e}^p = (\mathbf{1} + \mathbf{H}^p)^{-1} (\mathbf{1} + \boldsymbol{\chi}^p) - \mathbf{1} \simeq \boldsymbol{\chi}^p - \mathbf{H}^p, \text{ with } \boldsymbol{\chi}^p = \widehat{\boldsymbol{\chi}}^p - \mathbf{1} \quad (28)$$

When linearized, the state laws (19) become:

$$\boldsymbol{\sigma} = \rho \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}^e}, \mathbf{s} = \rho \frac{\partial \psi}{\partial \mathbf{e}^p}, \mathbf{M} = \rho \frac{\partial \psi}{\partial \mathbf{K}} \quad (29)$$

The evolution equations read then:

$$\dot{\boldsymbol{\varepsilon}}^p = \frac{\partial \Omega}{\partial (\boldsymbol{\sigma} + \mathbf{s})}, \dot{\boldsymbol{\alpha}} = -\frac{\partial \Omega}{\partial R} \quad (30)$$

The most simple case of a quadratic free energy potential is first considered:

$$\rho \psi(\boldsymbol{\varepsilon}^e, \mathbf{e}^p, \mathbf{K}) = \frac{1}{2} \boldsymbol{\varepsilon}^e : \mathbb{C} : \boldsymbol{\varepsilon}^e + \frac{1}{2} H_\chi \mathbf{e}^p : \mathbf{e}^p + \frac{1}{2} \mathbf{A} \mathbf{K} : \mathbf{K} \quad (31)$$

The usual four-rank tensor of elastic moduli is denoted by \mathbb{C} . The higher-order moduli have been limited to only two additional parameters: H_χ (unit MPa) and A (unit MPa.mm²). Their essential impact on the prediction of size effects will be analyzed in the next section. It follows that:

$$\boldsymbol{\sigma} = \mathbb{C} : \boldsymbol{\varepsilon}^e, \mathbf{s} = H_\chi \mathbf{e}^p, \mathbf{M} = A \mathbf{K} \quad (32)$$

Large values of H_χ ensure that \mathbf{e}^p remains small so that $\widehat{\chi}^p$ remains close to \mathbf{H}^p and \mathbf{K} is close to the dislocation density tensor. The yield condition for each slip system becomes:

$$f^s = |\tau^s - x^s| - \tau_c^s \quad (33)$$

with

$$x^s = -\mathbf{s} : \mathbf{P}^s = (\text{curl } \mathbf{M}) : \mathbf{P}^s = A (\text{curl } \text{curl } \chi^p) : \mathbf{P}^s \quad (34)$$

Comparison Between Micropolar and Micromorphic Crystal Plasticity

Experimental techniques like Electron Back-Scatter Diffraction (EBSD) provide the field of lattice orientation and, consequently, of lattice rotation \mathbf{R}^e during deformation. The rotation \mathbf{R}^e appears in the polar decomposition of the elastic deformation $\mathbf{F}^e = \mathbf{R}^e \mathbf{U}^e$, where \mathbf{U}^e is the lattice stretch tensor. Since

$$\boldsymbol{\alpha} = -\text{curl } \mathbf{F}^{e-1} = -\text{curl } (\mathbf{U}^{e-1} \cdot \mathbf{R}^{eT}) \quad (35)$$

the hypothesis of small elastic strain implies

$$\boldsymbol{\alpha} \simeq -\text{curl } \mathbf{R}^{eT} \quad (36)$$

This approximation also requires that the gradient of elastic strain is also small, which is not ensured even if the elastic strain is small. If, in addition, elastic rotations are small, we have:

$$\boldsymbol{\alpha} \simeq -\text{curl } (\mathbf{1} - \boldsymbol{\omega}^e) = \text{curl } \boldsymbol{\omega}^e \quad (37)$$

The small rotation axial vector is defined as:

$$\overset{\times_e}{\boldsymbol{\omega}} = -\frac{1}{2} \underset{\sim}{\boldsymbol{\varepsilon}} : \boldsymbol{\omega}^e, \boldsymbol{\omega}^e = -\underset{\sim}{\boldsymbol{\varepsilon}} \cdot \overset{\times_e}{\boldsymbol{\omega}} \quad (38)$$

or, in matrix notations:

$$[\omega^e] = \begin{bmatrix} 0 & \omega_{12}^e & -\omega_{31}^e \\ -\omega_{12}^e & 0 & \omega_{23}^e \\ \omega_{31}^e & -\omega_{23}^e & 0 \end{bmatrix} = \begin{bmatrix} 0 & -\overset{x_e}{\omega}_3 & \overset{x_e}{\omega}_2 \\ \overset{x_e}{\omega}_3 & 0 & -\overset{x_e}{\omega}_1 \\ -\overset{x_e}{\omega}_2 & \overset{x_e}{\omega}_1 & 0 \end{bmatrix} \tag{39}$$

The gradient of the lattice rotation field delivers the lattice curvature tensor. In the small deformation context, the gradient of the rotation tensor is represented by the gradient of the axial vector:

$$\kappa := \overset{x_e}{\omega} \tag{40}$$

One can establish a direct link between $\text{curl } \omega^e$ and the gradient of the axial vector associated with ω . For that purpose, the matrix form of $\text{curl } \omega^e$ is derived according to:

$$[\text{curl } \omega^e] = \begin{bmatrix} \omega_{12,3}^e + \omega_{31,2}^e & -\omega_{31,1}^e & -\omega_{12,1}^e \\ -\omega_{23,2}^e & \omega_{12,3}^e + \omega_{23,1}^e & -\omega_{12,2}^e \\ -\omega_{23,3}^e & -\omega_{31,3}^e & \omega_{23,1}^e + \omega_{31,2}^e \end{bmatrix} \tag{41}$$

or equivalently:

$$[\text{curl } \omega^e] = \begin{bmatrix} -\overset{x_e}{\omega}_{3,3} - \overset{x_e}{\omega}_{2,2} & \overset{x_e}{\omega}_{2,1} & \overset{x_e}{\omega}_{3,1} \\ \overset{x_e}{\omega}_{1,2} & -\overset{x_e}{\omega}_{3,2} - \overset{x_e}{\omega}_{1,1} & \overset{x_e}{\omega}_{3,2} \\ \overset{x_e}{\omega}_{1,3} & \overset{x_e}{\omega}_{2,3} & -\overset{x_e}{\omega}_{1,1} - \overset{x_e}{\omega}_{2,2} \end{bmatrix} \tag{42}$$

from which it becomes apparent that:

$$\alpha = \kappa^T - (\text{trace } \kappa) \mathbf{1}, \quad \kappa = \alpha^T - \frac{1}{2} (\text{trace } \alpha) \mathbf{1} \tag{43}$$

This is a remarkable relation linking, with the context of small elastic strains (and in fact of small gradients of elastic strain) and rotations, the dislocation density tensor to lattice curvature. It is known as Nye’s formula (Nye 1953).

As a conclusion, it appears that the Cosserat crystal plasticity model only considers the lattice curvature part contained in the full dislocation tensor. This seems to be a reasonable assumption. However, some significant differences can be found in the predictions of Cosserat vs. full micromorphic theory, as discussed in the reference Cordero et al. (2010a).

Size Effects in a Two-Phase Single-Crystal Laminate

A periodic two-phase single-crystal laminate under simple shear, whose unit cell is shown in Fig. 1, is considered, following Forest and Sedláček (2003a), Forest (2008), and Cordero et al. (2010a). This microstructure is composed of a hard elastic phase (h) and a soft elasto-plastic phase (s) where one slip system with slip direction normal to the interface between (h) and (s) is chosen. A mean simple glide $\bar{\gamma}$ is applied in the crystal slip direction \mathbf{l} of the phase (s). The displacement and microdeformation fields take the form:

$$u_1 = \bar{\gamma}x_2, u_2(x_1), u_3 = 0, \chi_{12}^p(x_1), \chi_{21}^p(x_1) \tag{44}$$

within the context of small deformation theory. It follows that:

$$\begin{aligned}
 [\mathbf{H}] &= \begin{bmatrix} 0 & \bar{\gamma} & 0 \\ u_{2,1} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
 [\mathbf{H}^p] &= \begin{bmatrix} 0 & \gamma & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad [\mathbf{H}^e] = \begin{bmatrix} 0 & \bar{\gamma} - \gamma & 0 \\ u_{2,1} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
 [\chi^p] &= \begin{bmatrix} 0 & \chi_{12}^p(x_1) & 0 \\ \chi_{21}^p(x_1) & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad [\text{curl } \chi^p] = \begin{bmatrix} 0 & 0 & -\chi_{12,1}^p \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}
 \end{aligned}$$

The resulting stress tensors are:

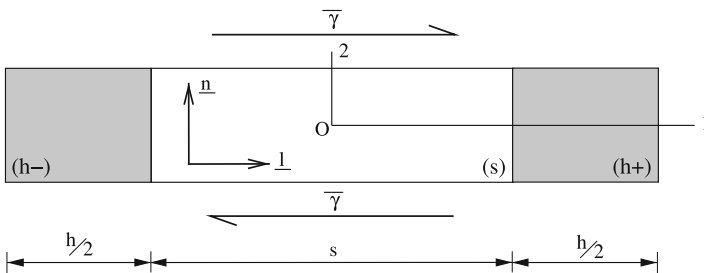


Fig. 1 Single slip in a periodic two-phase single-crystal laminate under simple shear: the gray phase (h) displays a purely linear elastic behavior, whereas the inelastic deformation of the white elasto-plastic phase (s) is controlled by a single-slip system (\mathbf{n}, \mathbf{l})

$$\begin{aligned}
[\sigma] &= \mu \begin{bmatrix} 0 & \bar{\gamma} - \gamma + u_{2,1} & 0 \\ \bar{\gamma} - \gamma + u_{2,1} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
[s] &= -H_\chi \begin{bmatrix} 0 & \gamma - \chi_{12}^p & 0 \\ -\chi_{21}^p & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
[M] &= \begin{bmatrix} 0 & 0 & -A\chi_{12,1}^p \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad [\text{curl } M] = \begin{bmatrix} 0 & -A\chi_{12,11}^p & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}
\end{aligned}$$

These forms of matrices are valid for both phases, except that $\gamma \equiv 0$ in the hard elastic phase. Each phase possesses its own material parameters, H_χ and A , the shear modulus, μ , being assumed for simplicity to be identical in both phases. The balance equation, $\mathbf{s} = -\text{curl } \mathbf{M}$, gives $\chi_{21}^p = 0$ and the plastic slip:

$$\gamma = \chi_{12}^p - \frac{A}{H_\chi} \chi_{12,11}^p. \quad (45)$$

In the soft phase, the plasticity criterion stipulates that:

$$\sigma_{12} + s_{12} = \tau_c + H\gamma_{\text{cum}}, \quad (46)$$

where H is a linear hardening modulus considered in this phase and γ_{cum} is the accumulated plastic slip as $\dot{\gamma}_{\text{cum}} = |\dot{\gamma}|$. The following analytical resolution is done for the first loading branch, under monotonic loading. The slip direction, \mathbf{l} , has been chosen such that $\gamma > 0$ for this first loading branch so that we have: $\gamma_{\text{cum}} = \gamma$. Considering Eqs. (45) and (46), we obtain the second-order differential equation for the microdeformation variable in the soft phase, χ_{12}^{ps} ,

$$\frac{1}{\omega^s} \chi_{12,11}^{ps} - \chi_{12}^{ps} = \frac{\tau_c - \sigma_{12}}{H}, \quad \text{with } \omega^s = \sqrt{\frac{H_\chi^s H}{A^s (H_\chi^s + H)}}. \quad (47)$$

where $1/\omega^s$ is the characteristic length of the soft phase for this boundary value problem. The force stress balance equation requires σ_{12} to be uniform. It follows that the nonhomogeneous part of the differential equation is constant and then the hyperbolic profile of χ_{12}^{ps} takes the form:

$$\chi_{12}^{ps} = C^s \cosh(\omega^s x) + D, \quad (48)$$

where C^s and D are constants to be determined. Symmetry conditions ($\chi_{12}^{ps}(-s/2) = \chi_{12}^{ps}(s/2)$) have been taken into account.

In the elastic phase, where the plastic slip vanishes, a hyperbolic profile of the microdeformation variable, χ_{12}^{ph} , is also obtained:

$$\chi_{12}^{ph} = C^h \cosh\left(\omega^h \left(x \pm \frac{s+h}{2}\right)\right), \text{ with } \omega^h = \sqrt{\frac{H\chi}{A^h}}, \quad (49)$$

where, again, C^h is a constant to be determined, and symmetry conditions have been taken into account. It is remarkable that the plastic microvariable, χ_{12}^{ph} , does not vanish in the elastic phase, close to the interfaces, although no plastic deformation takes place. This is due to the transmission of double traction. Such a transmission has been shown in Cordero et al. (2010a) to be essential for size effects to occur. This point will be discussed in section “[Size Effects in a Two-Phase Single-Crystal Laminate](#).” The linear constitutive equation for the double-stress tensor in (32) can be interpreted, for the elastic phase, as nonlocal elasticity. That is why the corresponding characteristic length, $1/\omega^h$, will be kept of the order of nanometers in the presented simulation.

The coefficients C^s , D , and C^h can be identified using the interface and periodicity conditions:

- Continuity of χ_{12}^p at $x = \pm s/2$:

$$C^s \cosh\left(\omega^s \frac{s}{2}\right) + D = C^h \cosh\left(\omega^h \frac{h}{2}\right). \quad (50)$$

- Continuity of the double traction, as given in Eq. (8), $m_{12} = -M_{13}$ at $x = \pm s/2$:

$$A^s \omega^s C^s \sinh\left(\omega^s \frac{s}{2}\right) = -A^h \omega^h C^h \sinh\left(\omega^h \frac{h}{2}\right). \quad (51)$$

- Periodicity of displacement component u_2 . We have the constant stress component:

$$\sigma_{12} = \mu (\bar{\gamma} - \gamma + u_{2,1}) \quad (52)$$

whose value is obtained from the plasticity criterion in the soft phase (Eq. 46):

$$\sigma_{12} = \tau_c + H\gamma_{cum} - A^s \chi_{12,11}^{ps}. \quad (53)$$

Still considering the first loading branch for which $\gamma_{cum} = \gamma$, it follows that:

$$u_{2,1}^s = \frac{\sigma_{12}}{\mu} - \bar{\gamma} + \gamma = \frac{\tau_c}{\mu} - \bar{\gamma} + \frac{A^s \omega^{s2} C^s}{H} \cosh(\omega^s x) + \frac{H + \mu}{\mu} D \quad (54)$$

in the soft phase and:

$$u_{2,1}^h = \frac{\sigma_{12}}{\mu} - \bar{\gamma} = \frac{\tau_c}{\mu} - \bar{\gamma} + \frac{H}{\mu} D \quad (55)$$

in the hard phase. The average on the whole structure,

$$\int_{-(s+h)/2}^{(s+h)/2} u_{2,1} dx = 0, \quad (56)$$

must vanish for periodicity reasons and gives

$$\left(\frac{\tau_c}{\mu} - \bar{\gamma} \right) (s+h) + \frac{2A^s \omega^s C^s}{H} \sinh\left(\omega^s \frac{s}{2}\right) + \frac{H(s+h) + \mu s}{\mu} D = 0 \quad (57)$$

The resolution of Eqs. (50), (51), and (57) gives:

$$C^s = \left(\frac{\tau_c}{\mu} - \bar{\gamma} \right) \left[\frac{A^s \omega^s \sinh\left(\omega^s \frac{s}{2}\right)}{s+h} \left(\frac{H(s+h) + \mu s}{\mu} \left(\frac{\coth\left(\omega^s \frac{s}{2}\right)}{A^s \omega^s} + \frac{\coth\left(\omega^h \frac{h}{2}\right)}{A^h \omega^h} \right) - \frac{2}{H} \right) \right]^{-1} \quad (58)$$

$$D = -A^s \omega^s C^s \sinh\left(\omega^s \frac{s}{2}\right) \left(\frac{\coth\left(\omega^s \frac{s}{2}\right)}{A^s \omega^s} + \frac{\coth\left(\omega^h \frac{h}{2}\right)}{A^h \omega^h} \right) \quad (59)$$

$$C^h = -C^s \frac{A^s \omega^s \sinh\left(\omega^s \frac{s}{2}\right)}{A^h \omega^h \sinh\left(\omega^h \frac{h}{2}\right)}. \quad (60)$$

Figure 2 shows the profiles of plastic microdeformation and double traction in the two-phase laminate for different sets of material parameters and for a fraction of soft phase (s), $f_s = 0.7$. These profiles clearly show the continuity of χ_{12}^p and m_{12} at the interfaces. The different shapes presented are obtained for various values of the modulus A^s , the other material parameters being fixed and given in Table 1. Varying A^s modifies the mismatch with respect to the modulus A^h of the phase (h). Without mismatch the profile of χ_{12}^p is smooth at interfaces, while stronger mismatches lead to sharper transitions between the phases. Varying A^s also changes the intrinsic length scale $1/\omega^s$ of the phase (s). When the intrinsic length scale is small compared to the size of the microstructure, the microdeformation gradient can develop inside the phase (s) which leads to a rounded profile of the plastic microdeformation χ_{12}^p and to a double traction m_{12} localized at the interfaces. When the intrinsic length scale increases, the value of the double traction also increases at the interfaces (or equivalently, when decreasing the microstructure length scale, $l = s + h$, for a fixed

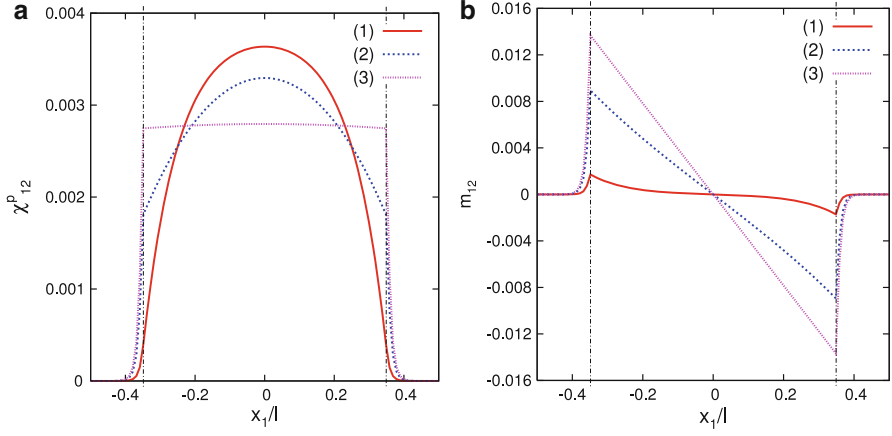


Fig. 2 Profiles of (a) plastic microdeformation χ_{12}^p and (b) double traction m_{12} in the two-phase microstructure with the *microcurl* model at 0.2% overall plastic strain obtained with the set of material parameters given in Table 1 and (1) with no mismatch between the moduli of the two phases, $A^h = A^s = 5.10^{-5}$ MPa.mm²; (2) with a stronger mismatch, $A^h = 5.10^{-5}$ MPa.mm² and $A^s = 1.10^{-3}$ MPa.mm²; and (3) $A^h = 5.10^{-5}$ MPa.mm² and $A^s = 5.10^{-2}$ MPa.mm². The associated intrinsic length scales, $1/\omega^s$, are, respectively, 100 nm, 449 nm, and 3.2 μ m. In all three cases, the fraction of soft phase $f_s = 0.7$ and the microstructure size is fixed, $l = 1 \mu$ m. The vertical lines indicate the position of interfaces

intrinsic length scale). When the intrinsic length scale becomes of the order of the size of the microstructure or even larger, the model starts to saturate so that χ_{12}^p becomes quasihomogeneous (flat profile) and the double traction is not localized anymore (linear profile). From Eq. (53) we derive the expression of the macroscopic stress tensor component, Σ_{12} , defined as the mean value of the stress component σ_{12} over the microstructure size, $l = (s + h)$:

$$\Sigma_{12} = \langle \sigma_{12} \rangle = \frac{1}{l} \int_{-\frac{l}{2}}^{\frac{l}{2}} \sigma_{12} dx = \tau_c + \frac{H}{f_s} \langle \gamma^{cum} \rangle - \frac{A^s}{f_s} \langle \chi_{12,11}^{ps} \rangle, \quad (61)$$

where brackets $\langle \rangle$ denote the average values over the microstructure unit cell. We obtain the mean plastic slip for the first loading branch from Eq. (45):

$$\langle \gamma \rangle = \left\langle \chi_{12}^{ps} - \frac{A^s}{H_\chi} \chi_{12,11}^{ps} \right\rangle = \frac{2A^s \omega^s C^s \sinh\left(\omega^s \frac{f_s l}{2}\right)}{Hl} + f_s D \quad (62)$$

where f_s is the fraction of soft phase. From this we obtain alternative expressions of C^s and D as functions of $\langle \gamma \rangle$,

Table 1 Set of material parameters used in the simulations. The intrinsic length scales, defined as $1/\omega^{h,s}$, induced by these parameters is of the order of 10 nm for the elastic phase (h) and 500 nm for the plastic phase (s)

	μ [MPa]	τ_c [MPa]	H [MPa]	H_k [MPa]	A [MPa.mm ²]
Phase (s)	35,000	40	5000	500,000	1.10^{-3}
Phase (h)	35,000	–	–	500,000	5.10^{-5}

$$C^s = -\langle \gamma \rangle$$

$$\left[A^s \omega^s \sinh \left(\omega^s \frac{f_s l}{2} \right) \left(f_s \left(\frac{\coth \left(\omega^s \frac{f_s l}{2} \right)}{A^s \omega^s} + \frac{\coth \left(\omega^h \frac{(1-f_s)l}{2} \right)}{A^h \omega^h} \right) - \frac{2}{Hl} \right) \right]^{-1} \quad (63)$$

$$D = \langle \gamma \rangle \left[f_s - \frac{2}{Hl} \left(\frac{\coth \left(\omega^s \frac{f_s l}{2} \right)}{A^s \omega^s} + \frac{\coth \left(\omega^h \frac{(1-f_s)l}{2} \right)}{A^h \omega^h} \right)^{-1} \right]^{-1} \quad (64)$$

which contain contributions from both the back stress and the isotropic hardening. The macroscopic stress takes the form:

$$\sum_{12} = \tau_c + HD. \quad (65)$$

The hardening produced by the model is a combination of the kinematic hardening arising from the higher-order back stress component and the linear isotropic hardening introduced in (46). Its modulus, H^{tot} , is size-dependent and is obtained using Eqs. (64) and (65):

$$H^{tot} = H \left[f_s - \frac{2}{Hl} \left(\frac{\coth \left(\omega^s \frac{f_s l}{2} \right)}{A^s \omega^s} + \frac{\coth \left(\omega^h \frac{(1-f_s)l}{2} \right)}{A^h \omega^h} \right)^{-1} \right]^{-1} \quad (66)$$

One cycle of deformation $\bar{\gamma}$ has been considered to illustrate the kinematic hardening effects. In the absence of gradient effects, only isotropic hardening is visible. The *microcurl* model leads to an additional kinematic hardening component. When the size of the elasto-plastic phase (s) becomes large compared to the intrinsic length scale $1/\omega_s$, strain gradient effect is small, and the kinematic hardening arising from the *microcurl* model tends to vanish. Then the model reduces to conventional crystal plasticity theory, and the limit of the 0.2% macroscopic flow stress is:

$$\lim_{l \rightarrow \infty} \Sigma_{12|0.2} = \tau_c + \frac{H}{f_s} \langle \gamma^{cum} \rangle. \quad (67)$$

In contrast, the maximum extra stress, $\Delta\Sigma$, predicted by the model at small microstructure sizes can be computed as:

$$\Delta\Sigma = \lim_{l \rightarrow 0} \Sigma_{12}(\langle \gamma \rangle) - \lim_{l \rightarrow \infty} \Sigma_{12|0.2} = \frac{1 - f_s}{f_s} H_\chi \langle \gamma \rangle. \quad (68)$$

Figure 3 presents the predicted evolution of the macroscopic flow stress $\Sigma_{12|0.2}$ at 0.2% plastic strain (obtained by setting $\langle \gamma \rangle = 0.002$) as a function of the microstructure length scale l in a log–log diagram. This evolution is plotted using the material parameters given in Table 1 and for various values of the coupling modulus, $H_\chi^s = H_\chi^h = H_\chi$. The four lower curves are obtained for finite values of the modulus H_χ ; they exhibit a tanh shape with saturation for large ($l > 10^{-2}$ mm) and small ($l < 10^{-5}$ mm) values of l . These saturations can be characterized by the limit given in Eq. (67) and the maximum extra stress, $\Delta\Sigma$, given in Eq. (68), respectively. A transition domain with strong size dependence is observed between these two plateaus. The limits and the maximum extra stress, the position of the transition zone, and the scaling law exponent in the size-dependent domain (slope in the log–log diagram) are directly related to the material parameters used in the model. In fact, the position of the size-dependent domain is controlled by the moduli $A^{h,s}$ (not illustrated here), while the maximum extra stress and the scaling law exponent are both controlled by the modulus H_χ , both increasing for higher values of H_χ as suggested by Fig. 3.

When H_χ is very small, we can deduce from Eq. (68) that $\Delta\Sigma$ vanishes, and consequently the scaling law exponent will tend to 0. The upper curve is obtained for $H_\chi \rightarrow \infty$; it no longer exhibits a tanh shape as no saturation occurs for small values of l , the limit $\Delta\Sigma \rightarrow \infty$ follows. This limit case will be described in the next subsection; it will be shown that in that case, a scaling law exponent of -2 is reached. Finally, the *microcurl* model can produce scaling law exponents ranging from 0 to -2 .

Strain Gradient Plasticity as a Limit Case

In the proposed *microcurl* model, the modulus H_χ introduces a coupling between micro and macro variables. A high value of H_χ forces the plastic microdeformation χ^p to remain as close as possible to the macroplastic deformation \mathbf{H}^p . Consequently, it enforces the condition that \mathbf{K} coincides with the dislocation density tensor. In this case, the *microcurl* model degenerates into the strain gradient plasticity model by Gurtin (2002). When applied to the laminate microstructure, the strain gradient plasticity model leads to the indeterminacy of the double-traction vector at the interfaces, due to the fact that no strain gradient effect occurs in the elastic phase; see Cordero et al. (2010a). The *microcurl* model can then be used to derive the missing

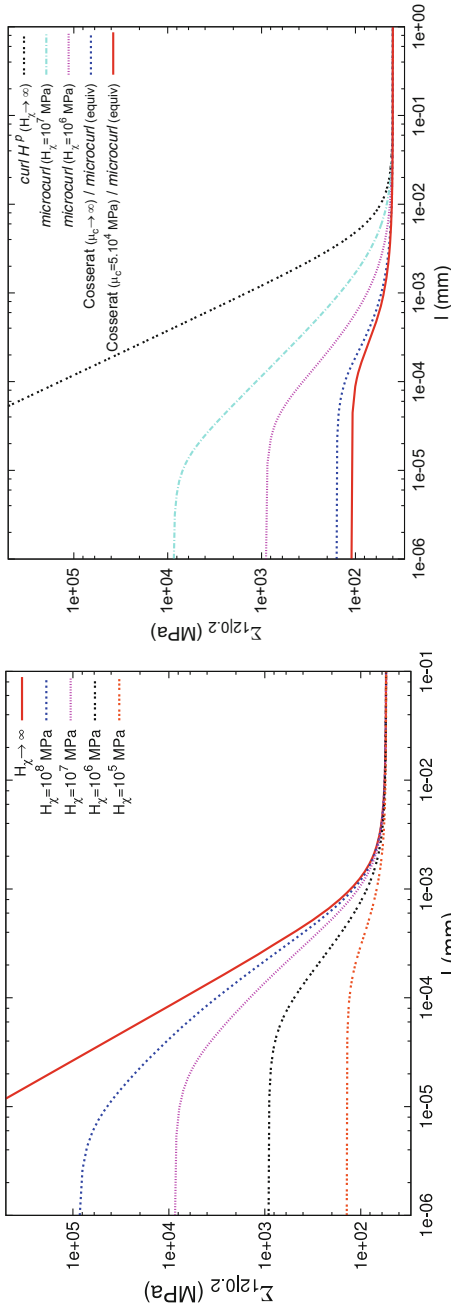


Fig. 3 Evolution of the macroscopic flow stress $\Sigma_{12|0.2}$ at 0.2% plastic strain as a function of the microstructure length scale l , plotted for different coupling moduli, H_χ ($= H_\chi^s = H_\chi^b$), on the left. The other material parameters are given in Table 1 and $f_s = 0.7$. On the right, comparison between the micropolar and micromorphic models

interface condition to be applied at the interface, by means of a limit process in the previous solution of the boundary value problem.

The limit $H_\chi \rightarrow \infty$ of the *microcurl* model can be used to determine the value of the double traction to be imposed at the interface, as follows:

$$\begin{aligned} \lim_{H_\chi \rightarrow \infty} m_{12}(s/2) &= \lim_{H_\chi \rightarrow \infty} A^s \chi_{12,1}^{ps}(s/2) \\ &= \lim_{H_\chi \rightarrow \infty} A^s \omega^s \sinh\left(\omega^s \frac{f_s l}{2}\right) C^s \\ &= \lim_{H_\chi \rightarrow \infty} \langle \gamma \rangle \left[\frac{2}{Hl} - f_s \left(\frac{\coth\left(\omega^s \frac{f_s l}{2}\right)}{A^s \omega^s} + \frac{\coth\left(\omega^h \frac{(1-f_s)l}{2}\right)}{A^h \omega^h} \right) \right]^{-1} \end{aligned}$$

Since $H_\chi \rightarrow \infty$, $1/\omega^h \rightarrow 0$ and $\coth(\omega^h h/2) \rightarrow 1$. Moreover, $\omega_\infty^s := \omega^s \rightarrow \sqrt{H/A^s}$.

Consequently,

$$\lim_{H_\chi \rightarrow \infty} m_{12}(s/2) = \langle \gamma \rangle \left[\frac{2}{Hl} - f_s \frac{\coth\left(\omega_\infty^s \frac{f_s l}{2}\right)}{A^s \omega_\infty^s} \right]^{-1} \quad (69)$$

Accordingly, the double traction is found to depend on the mean plastic slip. The characteristic length in the soft phase for the strain gradient plasticity model is found to be related to the ratio between the hardening modulus and the higher-order modulus, A^s . The limiting process can also be used to predict the response of the strain gradient plasticity model in the size effect zone. For that purpose, let us consider the limit of $\Sigma_{12|0,2}$, when H_χ goes to infinity. Indeed, when H_χ tends to infinity, the expression of D in Eq. (64) can be simplified. We consider sizes of the microstructures in the size effect zone, i.e., intermediate values of l . Since H_χ is very high, the term $\tanh(\omega^h(1-f_s)l/2)$ tends to 1. Considering that l is small enough, the term $l(\tanh(\omega_s f_s l/2))$ can be approximated by its Taylor expansion at the order 2, which leads to D of the form:

$$D \approx \frac{al + b}{cl^2 + dl + e} \quad (70)$$

where

$$a = \frac{\langle \gamma \rangle f_s}{2\sqrt{H_\chi}}, \quad b = \langle \gamma \rangle f_s A^h \left(1 + \frac{H}{H_\chi} \right) \quad (71)$$

$$c = -\frac{f_s^3 H \sqrt{A^h}}{12}, \quad d = \frac{f_s^2 H}{2\sqrt{H_\chi}}, \quad e = -\frac{f_s \sqrt{A^h} H}{H_\chi} \quad (72)$$

The terms a , d , and e tend to 0 when $H_\chi \rightarrow \infty$, so that:

$$D \approx \frac{12A_s \langle \gamma \rangle}{f_s^3 H l^2} \quad (73)$$

and for the macroscopic stress:

$$\Sigma_{12} \approx \tau_c + \frac{12A_s \langle \gamma \rangle}{f_s^3 l^2} \quad (74)$$

This expression indicates a l^{-2} scaling law for the strain gradient plasticity model. This scaling law differs from the Hall–Petch relation, $l^{-1/2}$, typical for grain size effects, and from the Orowan relation, l^{-1} , valid for precipitate size effects.

On the right Fig. 3 shows a comparison between the micropolar and micromorphic responses in the case of the two-phase laminate under shear. A saturation of the stress level is found for increasing values of the penalty modulus H_χ in the micropolar model at small scales. In contrast, the micromorphic response converges toward that of the strain gradient plasticity model and displays no limit at small scales. This is a fundamental difference between the lattice curvature-based and the dislocation density tensor-based models; see Cordero et al. (2010a).

Free Energy Potentials for Micromorphic Crystal Plasticity

The previous example showing that a simple quadratic potential with respect to the dislocation density tensor does not provide the satisfactory scaling law for the plastic behavior of the channel is an incentive for developing more physical constitutive laws for strain gradient plasticity. Such an attempt is presented in this section along the lines of Wulfinghoff et al. (2015).

Physically, the introduction of additional energy density terms may be motivated by the incomplete nature of the continuum theory. Clearly, the continuum description does not contain the full information on the discrete dislocation microstructure. In particular, single dislocations are not resolved. Instead, the continuum representation may be interpreted as a smoothed version of the real system, where information is lost deliberately. There is no reason to assume that the total elastic energy of the continuum representation coincides with the elastic energy of the real system including discrete dislocations. This is due to the loss of information as a result of the smoothing procedure (Mesarovic et al. 2010). Additional energy terms in gradient plasticity may therefore be interpreted as an attempt to partially compensate the error in the continuum elastic energy. This is done by taking into account available kinematical information on the dislocation microstructure as additional argument of the energy.

The optimal form of the energy is subject of current research. Most applications are based on a pragmatic quadratic approach (e.g., Cordero et al. 2012; Reddy et al.

2012; Wulfinghoff and Böhlke 2012; Miehe et al. 2014; Wulfinghoff et al. 2013a, b; Mesarovic et al. 2015).

Instead, a more reasonable approach seems to be based on a variable internal length scale as a function of the dislocation state (Groma et al. 2003; Mesarovic et al. 2010). The quadratic form was recently shown to provide physically unrealistic scaling in the size-dependent response of laminate microstructures under shear (Cordero et al. 2010b; Forest and Guéinichault 2013). Since quadratic forms are unusual in the classical dislocation theory, alternative free energy potentials were proposed in the past 10 years. Rank-one energies that are linear with respect to the GND densities have been shown to lead to a size-dependent yield stress in certain situations. Additionally motivated by line tension (and more elaborate) arguments, they are used by several authors (Ortiz and Repetto 1999; Conti and Ortiz 2005; Ohno and Okumura 2007; Kametani et al. 2012; Hurtado and Ortiz 2013).

Asymptotic methods can be used to derive alternative effective potential for distributions of edge dislocations. The asymptotic derivation of a logarithmic potential by De Luca et al. (2012) accounts for line tension effects at the macroscopic scale. Systematic derivations of back stress distributions were derived in Geers et al. (2013) by means of asymptotic methods.

The choice of a logarithmic energy is inspired by the statistical theory of dislocations of Groma et al. (2003, 2007) and Berdichevsky (2006a,b). Here, the internal length scale of the back stress is determined by the dislocation microstructure (see also Svendsen and Bargmann 2010; Forest and Guéinichault 2013). In the latter reference, the rank-one and logarithmic formulations were applied to strain gradient plasticity theories involving the full dislocation density tensor instead of the individual GND densities.

Formulation of Two Free Energy Potentials

It is assumed that the volumetric stored energy density has the form:

$$\rho\psi = W = W_e + W_g + W_h, \quad (75)$$

with $W_e = (\varepsilon - \varepsilon^p) : \mathbb{C} : (\varepsilon - \varepsilon^p)/2$. The expressions W_h and W_g are assumed to be functions of internal (hardening) variables α and the dislocation density tensor \mathbf{A} , respectively.

Size-independent isotropic hardening is accounted for by W_h , while W_g models size effects.

The following defect energies are investigated:

$$W_g^l = cGb \|\mathbf{A}\|, \quad W_g^{\ln} = c_0 \|\mathbf{A}\| \ln \frac{\|\mathbf{A}\|}{A_0}, \quad (76)$$

where c is a constant of order unity, G is the macroscopic shear modulus, b is the Burgers vector, A_0 is a constant, and c_0 is given by:

$$c_0 = \frac{Gb\beta}{2\pi(1-\nu)}, \quad (77)$$

where ν is Poisson's ratio and β is of order unity. The Euclidean norm of the dislocation density tensor is defined as: $\|\mathbf{A}\| = \sqrt{\mathbf{A} \cdot \mathbf{A}}$.

The rank-one energy W_g^1 can be motivated by simple line tension arguments; see Ortiz and Repetto (1999) and Hurtado and Ortiz (2012, 2013).

The logarithmic energy W_g^{\ln} (Eq. (76)) is motivated by the form of the associated back stress (Forest and Guéninchault 2013). It turns out that the approach W_g^{\ln} leads to a back stress which is formally close to the one derived in the statistical theory of Groma et al. (2003), given in 1D by

$$\frac{Gc_1}{2\pi(1-\nu)\rho} \partial_{x_1}^2 \gamma \quad (78)$$

for a single-slip situation with slip direction e_1 . Here, ρ denotes the total dislocation density. In the two-dimensional single-slip regime, the back stress involves the Laplacian of the plastic slip, as postulated by Aifantis (1987). However, the internal length scale is not interpreted as a material constant but determined by the dislocation microstructure, if W_g^{\ln} is applied.

The stresses σ and \mathbf{M} are assumed to be energetic, i.e.,

$$\sigma = \frac{\partial W}{\partial \boldsymbol{\epsilon}^e}, \quad \mathbf{M} = \frac{\partial W}{\partial \mathbf{A}} \quad (79)$$

If the stored energy is not differentiable at $\mathbf{A} = \mathbf{0}$, the symbol ∂ in Eq. (79)₂ is interpreted as a subdifferential operator (see, e.g., Han and Reddy 2013), i.e.,

$$\mathbf{M}|_{\mathbf{A}=\mathbf{0}} \in \{ \mathbf{M} : W_g(\mathbf{A}) - \mathbf{M} \cdot \mathbf{A} \geq 0 \quad \forall \mathbf{A} \}. \quad (80)$$

This can be interpreted as follows. If the stress \mathbf{M} is applied at a material point, \mathbf{A} will take a value which minimizes the expression $W_g(\mathbf{A}) - \mathbf{M} \cdot \mathbf{A}$. For small values of \mathbf{M} , the minimum is given by $\mathbf{A} = \mathbf{0}$. However, for sufficiently large values of \mathbf{M} , the value of \mathbf{A} can be determined from the stationarity condition $\mathbf{M} = \partial_{\alpha} W_g$.

For example, if W_g is given by $W_g^1 = cGb \|\mathbf{A}\|$, it follows that $\mathbf{A} = \mathbf{0}$ if

$$\mathbf{M} : \mathbf{A} \leq W_g^1(\mathbf{A}) = cGb \|\mathbf{A}\| \quad \forall \mathbf{A} \iff \mathbf{M} : \mathbf{A} \leq \|\mathbf{M}\| \|\mathbf{A}\| \leq cGb \|\mathbf{A}\| \quad \forall \mathbf{A}. \quad (81)$$

Hence, it is found that:

$$\left\{ \begin{array}{l} \mathbf{M} \in \{ \mathbf{M} : \varphi(\mathbf{M}) \leq 0 \}, \text{ if } \mathbf{A} = \mathbf{0} \\ \mathbf{M} = cGb \frac{\mathbf{A}}{\|\mathbf{A}\|}, \text{ else.} \end{array} \right. \quad (82)$$

with $\varphi(\mathbf{M}) = \|\mathbf{M}\| - cGb$.

Note that the generalized stress \mathbf{M} can be computed uniquely from \mathbf{A} only if $\mathbf{A} \neq 0$. This makes analytical solutions as well as the numerical implementation difficult. Possible numerical strategies concerning this problem are discussed in Kametani et al. (2012) as well as Hurtado and Ortiz (2013).

Application to the Shearing of the Periodic Laminate

In this section, the two new potentials are applied to the elasto-plastic laminate microstructure already considered in section “Size Effects in a Two-Phase Single-Crystal Laminate”; see Fig. 1. The two promising candidates of the defect energy function W_g are investigated concerning their effect on the overall size effects as well as the dislocation pileup structures building up at impenetrable boundaries. The dislocation density tensor can be expressed in terms of the edge density $\rho_{\perp} = -\partial_{x_1}\gamma$

$$\mathbf{A} = -\rho_{\perp} \mathbf{e}_1 \otimes \mathbf{e}_3. \quad (83)$$

The quantity ρ_{\perp} represents the total Burgers vector amount per unit area of edge dislocations. Note that its unit (μm^{-1}) differs from the unit of the total line length per unit volume ρ , given by μm^{-2} .

Assuming the defect energy W_g to be a function of $\|\mathbf{A}\|$, the generalized stress \mathbf{M} reads

$$\mathbf{M} = \partial_A W_g = \partial_{\|\mathbf{A}\|} W_g \frac{\mathbf{A}}{\|\mathbf{A}\|} = M(x_1) \mathbf{e}_1 \otimes \mathbf{e}_3. \quad (84)$$

From the balance Eq. (7)₂, it follows that

$$s_{12} - M' = 0. \quad (85)$$

Throughout this section, the isotropic hardening contribution will be neglected, i.e., $W_h = 0$.

Rank-One Defect Energy

For the laminate, the following energy is adopted:

$$W_g^1 = cGb \|\mathbf{A}\| = cGb |\rho_{\perp}|, \quad (86)$$

where c is of order unity (Ortiz and Repetto 1999). According to Eq. (84), the generalized stress \mathbf{M} reads:

$$\mathbf{M} = -\frac{\rho_{\perp}}{|\rho_{\perp}|} cGb = -\text{sign } \rho_{\perp} cGb, \text{ if } |\rho_{\perp}| > 0 \\ |M| \leq cGb, \text{ if } |\rho_{\perp}| = 0 \quad (87)$$

where the second line follows from Eq. (82).

Subsequently, a monotonic shear deformation in the positive direction is prescribed such that the following relations hold in the soft phase: $\tau^{\text{eff}} \geq \tau^C$, $\dot{\gamma} \geq 0$.

The principle of virtual power is written for the real field on the laminate unit cell V :

$$\int_V \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} + s : \dot{\mathbf{H}}^p + \mathbf{M} : \text{curl } \dot{\mathbf{H}}^p dV = \int_{\partial V} \mathbf{t} \cdot \dot{\mathbf{u}} dS + \mathbf{m} : \dot{\mathbf{H}}^p dS \quad (88)$$

The last term of the right-hand side vanishes due to the fact that \mathbf{H}^p is periodic, whereas \mathbf{m} is antiperiodic. The first term of the right-hand side coincides with the first term in the left-hand side, as can be shown by means of the Gauss theorem. As a result we obtain,

$$\int_{\partial V} s : \dot{\mathbf{H}}^p + \mathbf{M} : \text{curl } \dot{\mathbf{H}}^p dV = 0 \quad (89)$$

For the laminate under single slip, this gives:

$$\int_{-s/2}^{s/2+h} s_{12} \dot{\gamma} + (-cGb \text{sign } \gamma_{,1}) (-\dot{\gamma}_{,1}) dx_1 = 0 \quad (90)$$

Under monotonic loading, $\text{sign } \gamma_{,1} = \text{sign } \dot{\gamma}_{,1}$, so that

$$\int_{-s/2}^{s/2+h} s_{12} \dot{\gamma} + cGb |\dot{\gamma}_{,1}| dx_1 = 0 \quad (91)$$

According to the Schmid law, $\sigma_{12} + s_{12} = \tau^{rmC}$ where the fields σ_{12} , s_{12} are uniform. The solution is such that the plastic slip field $\gamma(x_1)$ is uniform in $] -s/2, s/2[$ at each instant. So does $\dot{\gamma}(x_1) = \dot{\gamma}(0)$ in $] -s/2, s/2[$. It vanishes in $]s/2, s/2 + h[$. It is therefore discontinuous at $\pm s/2$. As a result, the derivative of the plastic slip rate is the sum of two Dirac functions:

$$\dot{\gamma}_{,1}(x_1) = \dot{\gamma}(0) \left(\delta \left(x_1 + \frac{s}{2} \right) - \delta \left(x_1 - \frac{s}{2} \right) \right) \quad (92)$$

The integration of these Dirac functions (in fact the absolute values due to (91)) on the interval $[-s/2, s/2 + h]$ finally gives

$$s (\tau^{rmC} c - \sigma_{12}) \dot{\gamma}(0) + 2cGb \dot{\gamma}(0) = 0 \quad (93)$$

The scaling law follows:

$$\sigma_{12} = \tau^C + \frac{2cGb}{s}. \quad (94)$$

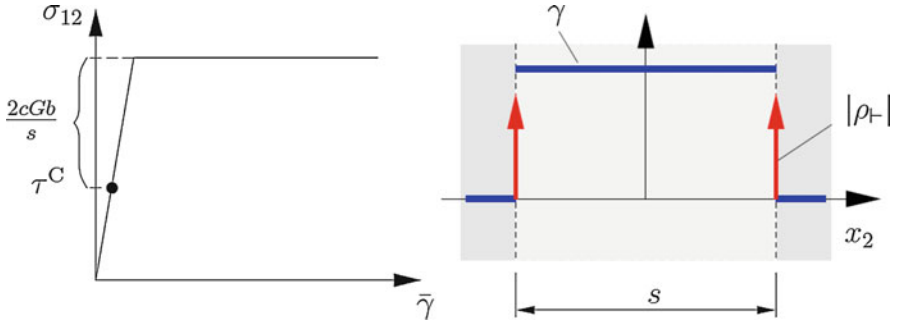


Fig. 4 Macroscopic shear stress–strain curve for the rank-one energy. The increase of the overall yield point scales inversely with the size of the soft phase, after Wulfinghoff et al. (2015)

This equation holds in the plastic regime. Clearly, the application of the rank-one energy increases the macroscopic yield stress by $2cGb/s$, i.e., the increase scales inversely with the size of the soft phase (see Fig. 4). The same scaling behavior has been found by Ohno and Okumura (2007) for a spherical grain, also using a rank-one energy. The authors concentrated on the overall mechanical response without having to compute the fields inside of the grains. As illustrated in Fig. 4, the dislocations localize in dislocation walls at the elasto-plastic interface.

For the material parameters of aluminum ($G = 26.12$ GPa and $b = 0.286$ nm) and $c = 1$, the size effect becomes important when the system size is below $10 \mu\text{m}$. The plastic shear strain is constant in the bulk, i.e., the dislocations form singularities (walls) at the boundaries. The back stress is constant (w.r.t. space) in the bulk. During the first period, it increases and thereby impedes any plastic deformation. Therefore, the overall deformation is purely elastic during this period. At a certain point, the plastic deformation starts, and the back stress remains constant afterward. Its value is given by $2cGb/s$.

Logarithmic Energy

This section investigates the following defect energy:

$$W_g^{\text{ln}} = c_0 \|\mathbf{A}\| \ln \frac{\|\mathbf{A}\|}{A_0}, \quad (95)$$

with the constant c_0 as defined in Eq. (77). The energy is motivated by the statistical theory of dislocations by Groma et al. (2003). The authors derive a back stress term which involves the second gradient of slip as postulated by Aifantis (1987). However, their theory involves an internal length scale which is given by $1/\sqrt{\rho}$, where ρ denotes the total dislocation density.

In pure metals, the geometrical characteristics of the microstructure are essentially determined by the dislocation arrangement. This is a strong argument for a (variable) internal length scale, which is determined by the available dislocation

field variables (instead of a constant length scale parameter; see also Forest and Sedláček (2003b) where this dependency is derived from a dislocation line tension model).

It is demonstrated subsequently that the approach (95) leads to a back stress which is similar to that of Groma et al. (2003). However, it should be mentioned that this energy is neither convex nor smooth with respect to the dislocation density tensor (a regularization will be discussed at a later stage).

For the laminate problem, the generalized stress M reads (see Eq. (84)):

$$M = -\text{sign } \rho_{\perp} c_0 \left(\ln \frac{|\rho_{\perp}|}{A_0} + 1 \right). \quad (96)$$

In this section, rate-independent plasticity will be considered based on the yield criterion:

$$f = |\tau^{\text{eff}}| - \tau^{\text{C}} \leq 0. \quad (97)$$

Here, the effective stress reads:

$$\tau^{\text{eff}} = (\boldsymbol{\sigma} + s) \cdot (\mathbf{l} \otimes \mathbf{n}) = \sigma_{12} + s_{12} \stackrel{(85)}{=} \sigma_{12} + M'. \quad (98)$$

With Eq. (96) and $M' = (\partial_{\rho_{\perp}} M) (\partial_{x_1} \rho_{\perp})$, it follows that

$$\tau^{\text{eff}} = \tau - \frac{c_0}{|\rho_{\perp}|} \partial_{x_1} \rho_{\perp} = \tau + \frac{G\beta}{2\pi(1-\nu)} \frac{b}{|\rho_{\perp}|} \partial_{x_1}^2 \gamma. \quad (99)$$

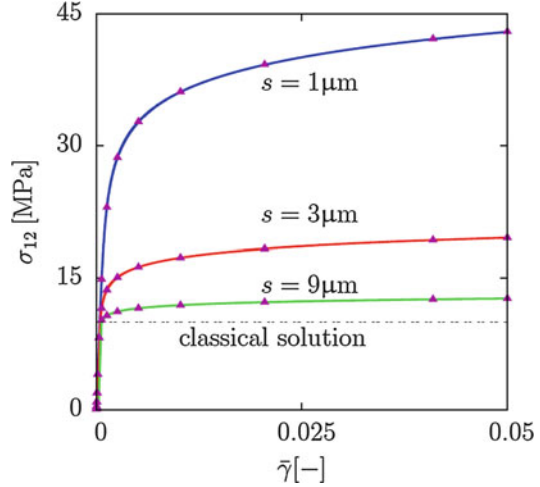
Here, the second term can be interpreted as a back stress. Note that the back stress involves no internal length scale parameter. Instead, the internal length scale, $\sqrt{b/\rho_{\perp}}$, is determined by the dislocation microstructure. In contrast to the back stress of Groma et al. (2003), the internal length scale is determined by the GND density ρ_{\perp} instead of the total density ρ . Hence, the influence of statistically stored dislocations (SSDs) is ignored for brevity. Therefore, a homogeneous initial GND density $|\rho_{\perp}| = A_0$ will be assumed to be given. In addition, it is assumed that the SSD density is equal or less than A_0 . The soft phase is assumed to be under plastic loading, with $\tau^{\text{eff}} = \tau^{\text{C}}$ in the soft phase. In this case,

$$M' \stackrel{(85)}{=} s_{12} = -(\sigma_{12} - \tau^{\text{C}}) = \text{const.} \Rightarrow M = -(\sigma_{12} - \tau^{\text{C}}) x_1, \quad (100)$$

where, again, the constant of integration vanishes due to the symmetry requirement $|M(-s/2)| = |M(s/2)|$. The plastic slip γ can be derived from the equality of Eqns. (96) and (100), which yields a differential equation for γ . The solution reads:

$$\gamma = \frac{A_0 L}{e} \left(\exp\left(\frac{s}{2L}\right) - \exp\left(-a \frac{x_1}{L}\right) \right) \text{ with } L = \frac{c_0}{\sigma_{12} - \tau^{\text{C}}}, \quad (101)$$

Fig. 5 Macroscopic stress–strain diagram for three different sizes. Analytical (*lines*) and regularized, numerical (*triangles*) solution for the logarithmic potential, after Wulfinghoff et al. (2015)



where the matching conditions $\gamma(-s/2) = \gamma(s/2) = 0$ have been exploited and where $e = \exp(1)$. The variable a is defined by $a = \text{sign } \gamma'$ and is assumed positive in $(-s/2, 0)$ and negative in $(0, s/2)$.

The macroscopic stress–strain relation follows:

$$\bar{\gamma} = \frac{A_0 L}{e(s+h)} \left(\exp\left(\frac{s}{2L}\right) (s-2L) + 2L \right) + \frac{\sigma_{12}}{G}. \quad (102)$$

The solution is evaluated for the following material parameters:

E [GPa]	ν	τ^c [MPa]	b [nm]	β	A_0/b [μm^{-2}]
70	0.34	10	0.286	1	1

A very thin hard phase with negligible width h is considered ($h/s = 10^{-6}$ for the analytical solution).

The macroscopic stress–strain curve (102) is illustrated in Fig. 5. A clear size effect is visible. Apparently, mainly the overall yield stress is affected. The hardening shows less size dependence. It is remarkable that the model provides a size-dependent yield stress and nonlinear kinematic hardening.

Since there is no distinct yield stress, the evaluation of the scaling behavior is based on the offset yield stress at 0.2% plastic strain. The offset yield stress as a function of the inverse of the size $1/s$ exhibits the same behavior as in the results obtained from the rank-one energy. It scales inversely with channel size; see Wulfinghoff et al. (2015).

Regularization of the Logarithmic Energy

The following regularization is introduced:

$$W_g = \begin{cases} \frac{1}{2} \frac{c_0}{b} l^2 \|A\|^2, & \|A\| < A_L \\ c_0 \|A\| \ln \frac{\|A\|}{A_0} + W_0, & \text{else.} \end{cases} \quad (103)$$

In the region of small GND densities, the energy is replaced by a quadratic potential. The internal length scale l , the transition density α_L , and the offset energy W_0 are chosen such that W_g , $\partial_{\|A\|} W_g$ and $\partial_{\|A\|}^2 W_g$ are continuous at the transition point $\|A\| = A_L$. As a result

$$A_L = A_0, \quad l^2 = \frac{b}{A_0}, \quad W_0 = \frac{c_0 A_0}{2}. \quad (104)$$

The regularized energy (103) is convex, normalized, and twice differentiable. The back stress for the laminate problem reads:

$$x = \begin{cases} -\frac{c_0}{A_0} \partial_{x_1}^2 \gamma, & |\rho_{\perp}| < A_0, \\ -\frac{c_0}{|\rho_{\perp}|} \partial_{x_1}^2 \gamma, & \text{else.} \end{cases} \quad (105)$$

Cyclic Behavior of the Laminate

The laminate is now submitted to one full cycle $\bar{\gamma} = \pm 0.05$. The hysteresis loops σ_{12} vs. $\bar{\gamma}$ for both rank-one and logarithmic potentials are represented in Fig. 6 for $s = 3 \mu\text{m}$. In the absence of isotropic hardening, the loops are stabilized after one full cycle. They are characterized by pure kinematic hardening. The influence of the back stress is clearly observable. The curves in Fig. 6 have been obtained numerically. One striking feature of the results is the nonconvexity of the obtained loops. According to the rank-one model, the first unloading stage is characterized by reverse plasticity at a constant negative shear stress. When $\bar{\gamma}$ goes through zero again, the overall shear stress experiences a jump of the same magnitude as computed analytically for monotonic loading in sections “[Rank-One Defect Energy](#)” and “[Logarithmic Energy](#).” The nonconvex loop obtained for the logarithmic potential is similar but smoother and displays smooth nonlinear kinematic hardening. A similar nonconvex hysteresis loop was obtained by Ohno and Okumura (2008) for the rank-one model.

The type of nonlinear kinematic hardening observed for both models can be identified with Asaro’s type KIII model, corresponding to a *first-in/last-out* sequence of dislocation motion (Asaro (1975)). It is considered by Asaro as the most perfect form of recovery of plastic memory. Such stress–strain loops display inflection points that are observed in some materials, see Asaro (1975) for a Nimonic alloy, but such observations have also been made in several Nickel based superalloys. It is usually attributed to substructural recovery on the microscale,

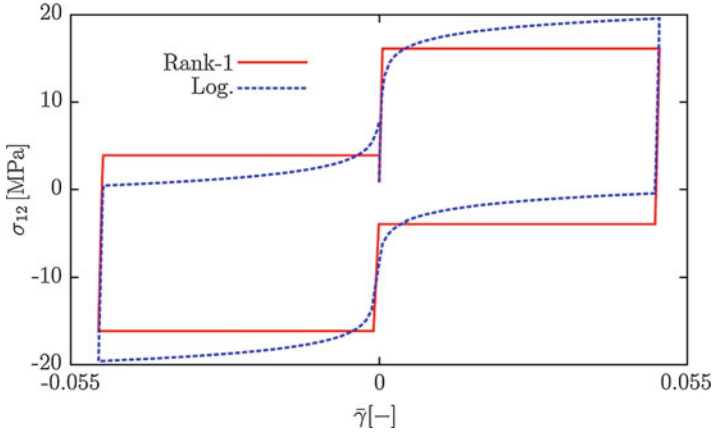


Fig. 6 Cyclic loading for $s = 3\mu\text{m}$

for instance pileup formation and destruction at γ' precipitates. In the present simple single-crystal model, it is the single active hardening mechanism induced by strain gradient plasticity and the presence of the hard phase in the laminate. It represents an accurate continuum description of dislocation piling-up and unpiling-up phenomena.

The experimental evidence of such nonconvex loops is illustrated in Figs. 7 and 8 in the case of polycrystalline Fe–Cr and Al–Cu–Mg alloys, respectively. The first loop in Fig. 7 (left) exhibits two inflection points, but the convexity is restored after a few cycles, and the usual shape with still a strong Bauschinger effect is retrieved in Fig. 7 (right). Figure 8 shows that the amount of plastic recoverability is controlled by the annealing degree of the dislocation microstructure. Further evidence of nonconvex loops in the cyclic behavior of FCC alloys can be found in the recent contribution by Proudhon et al. (2008) dealing with aluminum alloys. The common characteristics of these FCC alloys are that they all contain a population of nonshearable intragranular precipitates. This distribution of particles represents the first series of obstacles to be overcome by dislocations for the plasticity to start. The distance between precipitates presents a small scatter, and the average value is the characteristic length responsible for the size-dependent yield limit. This distance is comparable to the width s in our ideal laminate model. As illustrated by the TEM observations by Stoltz and Pelloux (1974, 1976), Taillard and Pineau (1982), and Proudhon et al. (2008), dislocation loops multiply around precipitates and can be destroyed after reverse loading unless the material is annealed before reversing the load, see Fig. 8, or unless the multiplication of forest dislocations or cross-slip effects limit the recoverability of cyclic plasticity. The effect has also been observed in nickel-base single-crystal superalloys for tension along $\langle 111 \rangle$; see Fig. 9. The simulations based on the logarithmic potential provide smooth loops that are closer to the experimental shapes. Our simulations deal with ideal single-

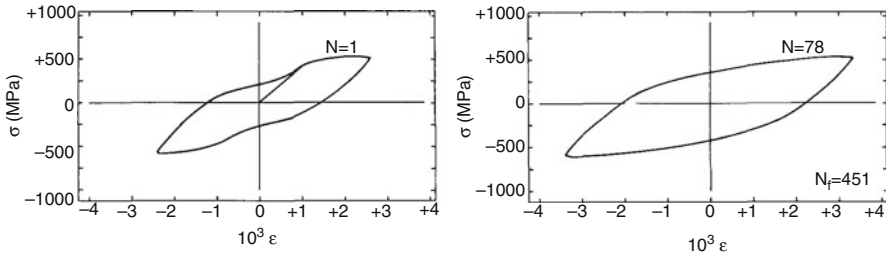


Fig. 7 Shape of the stress–strain hysteresis loop as a function of the number N of cycles for a Fe-19wt.%Cr alloy aged at 923 K for 72 h and mechanically tested at room temperature: $N = 1$ (left), $N = 48$ (right, N_f indicates the number of cycles to failure), after Taillard and Pineau (1982)

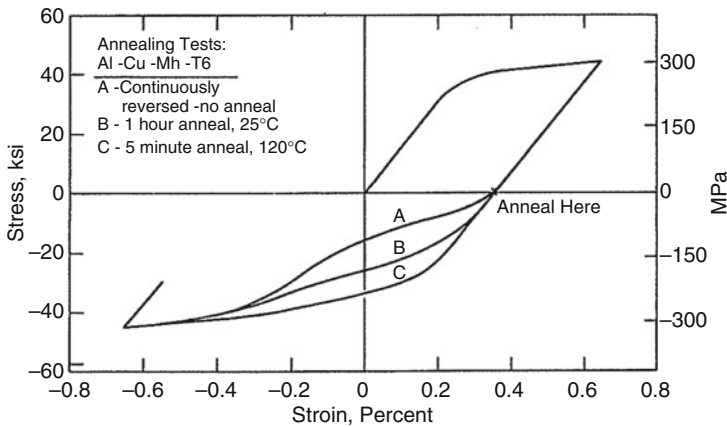


Fig. 8 Interrupted annealing hysteresis curves for an Al–Cu–Mg–T6 alloy tested at room temperature: A, continuously reverse load (no anneal); B, annealed 1 h/25 °C; and C, annealed 5 min/120 °C, after Stoltz and Pelloux (1976)

crystal laminates and simulations for polycrystals remain to be done. However, as shown by the two-dimensional strain gradient plasticity simulations performed by Ohno and Okumura (2008), based on the rank-one potential, the effect pertains for polycrystals. However, these authors did not recognize the physical reality of the simulated phenomena. Instead, they further developed the model to replace the rank-one energy potential by a dissipative formulation which restores the convexity of fatigue loops.

The two nonquadratic energy potentials represent continuum models of a discrete phenomenon which can be illustrated for a single dislocation source, as shown in Fig. 10. The cyclic response of a Frank–Read source, simulated by discrete dislocation dynamics (DDD) (Déprés et al. 2004; Chang et al. 2016), provides a nonconvex loop which is identical to the one predicted by the rank-one continuum model. This is related to the fact that an instability of the loop behavior is observed

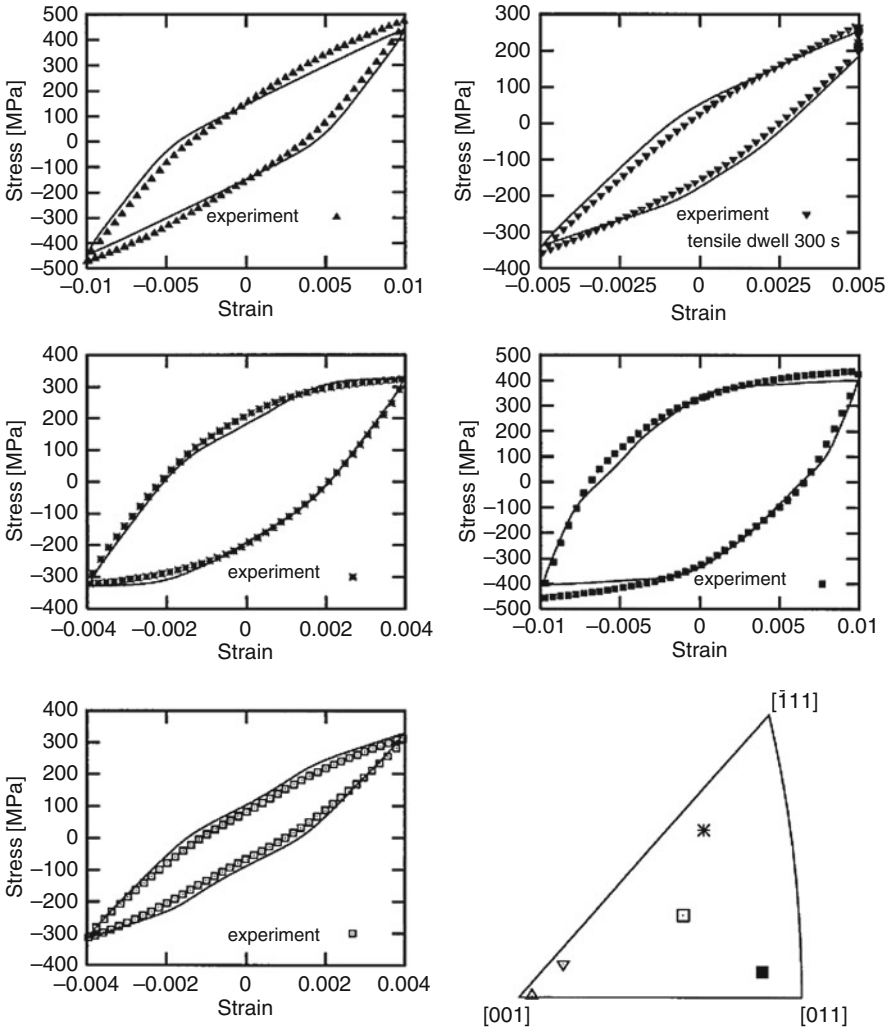


Fig. 9 Stabilized stress–strain loops for nickel-base single-crystal superalloy SC16 at 950 °C, experiment vs. simulation after Fedelich (2002)

for a critical stress that is inversely proportional to the length of the source. The scenario of dislocation source bowing and sudden propagation and multiplication can be reversed entirely in the absence of strong interaction with the dislocation forest and in the absence of cross-slip. This explains the concave shape of the stress–strain loop predicted by the DDD, which is accurately translated by the continuum model; see Fig. 6. Statistical effects of the collective behavior of dislocations finally destroy the recoverability of plastic deformation and the associated transmission of the single source behavior to the macroscopic response.

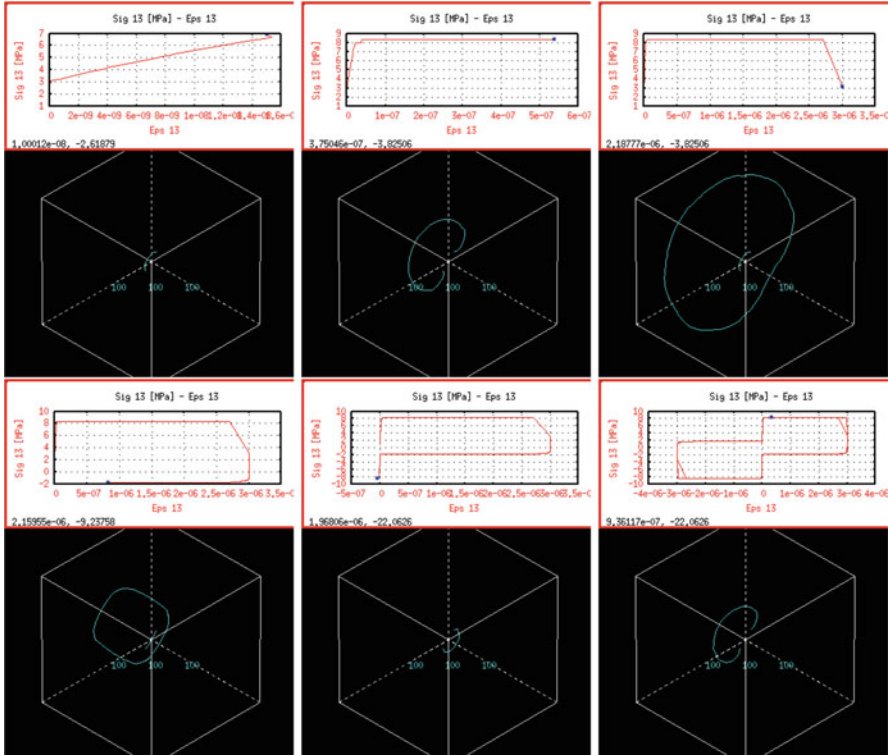


Fig. 10 Cyclic behavior of a single Frank–Read dislocation source simulated by discrete dislocation density. (Courtesy of Dr. M. Fivel)

Grain Size Effects in Polycrystals

The model is now applied to simulate the response of polycrystals and the effects of grain size.

The interface conditions at grain boundaries play a major role in the simulated size effects in the polycrystal behavior. No special interface law is considered in this work, although such physically motivated interface conditions exist in the literature; see Gurtin and Anand (2008). Instead, we consider the canonical interface conditions that arise from the formulation of the balance equations of the *microcurl* continuum model. These conditions are the continuity of displacement, \mathbf{u} , and the continuity of plastic microdeformation, χ^p . These conditions also include the continuity of the simple and double tractions, \mathbf{t} and \mathbf{M} , described in Eq. (8). Continuity of displacement excludes grain boundary cracking and sliding. Continuity of plastic microdeformation is reminiscent of the fact that dislocations generally do not cross grain boundaries, especially for such random grain boundaries. Note that in the *microcurl* model, only the kinematic degrees of freedom χ^p are continuous.

This is not the case of the plastic deformation, \mathbf{H}^p , which is treated here as an internal variable. However, due to the internal constraint discussed in section “[Model Formulation](#),” \mathbf{H}^p closely follows the plastic microdeformation so that it is quasi-continuous at grain boundaries when the penalty coefficient, H_χ , is high enough. Conversely, lower values of H_χ may allow slightly discontinuous plastic deformation, which may be tentatively interpreted as dislocation sinking inside grain boundaries. The continuity of the associated tractions expresses the transmission of classical and generalized internal forces from one grain to another through grain boundaries. Such continuum models are then able to mimic in that way the development of dislocation pileups at grain boundaries (Forest and Sedláček 2003a).

More elaborate grain boundary behavior laws are necessary to go beyond the three possible interface conditions readily available according to the *microcurl* model: vanishing microdeformation, continuous microdeformation, or vanishing microtractions at grain boundaries. They require proper account of transmission and absorption rules for dislocations at grain boundaries. A simple and efficient strategy was proposed for the formulation and finite element implementation of such interface constitutive laws by Wulfinghoff et al. (2013a). The reader is referred to the references quoted therein for more advanced grain boundary behavior laws.

Boundary Value Problem for Polycrystals

The size effects exhibited by the solution of the boundary value problem are linked to an intrinsic length scale, l_s , introduced through the generalized moduli H_χ and A of Eq. (32) and defined as:

$$l_s = \sqrt{\frac{A}{H_\chi}}. \quad (106)$$

This intrinsic length scale has to be consistent with the fact that plasticity effects occur at scales ranging from hundreds of nanometers to a few microns. In addition, as stated in section “[Model Formulation](#),” the coupling modulus, H_χ , has to be chosen high enough to ensure that χ^p and \mathbf{H}^p are close. These requirements are guidelines for the choice of relevant generalized moduli H_χ and A . The sets of material parameters used in this section are chosen in that way.

The finite element simulations have been made on periodic 2D meshes of periodic polycrystalline aggregates generated by a method based on Voronoi tessellations (Fig. 11a, b). Quadratic isoparametric finite elements with reduced integration are used. The random distribution of the grain centers has been controlled so that their sizes are sensibly the same, around the mean grain size, d . A random orientation is assigned to each grain, and two slip systems are taken into account for simplicity. In 2D, the plastic behavior of FCC crystals can be simulated with 2D planar double slip by considering two effective slip systems separated by an angle of 2ϕ (Asaro 1983; Bennett and McDowell 2003). Figure 11c describes the geometry.

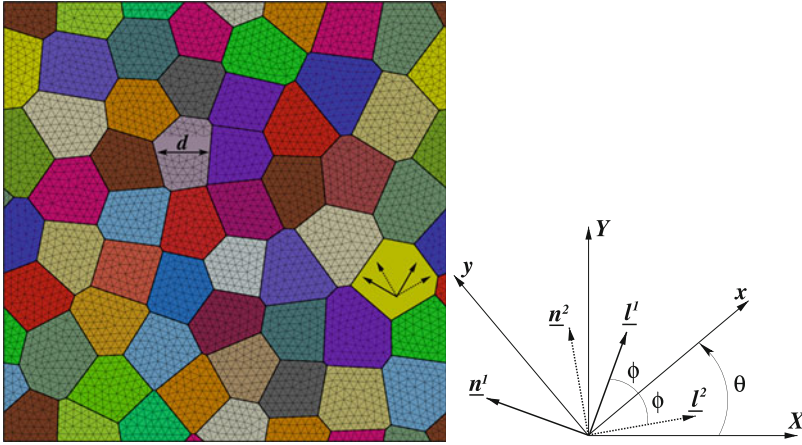


Fig. 11 Periodic meshes of the 2D periodic aggregates used in the finite element simulations including 52 grains. Two slip systems are taken into account in each randomly oriented grain. Various mean grain sizes, d , ranging from tens of nanometers to hundreds of microns, are investigated. On the right, description of the two effective slip systems for 2D planar double slip

The slip system pair is oriented by the angle θ , which is the grain orientation randomly fixed for each grain. For a FCC crystal $\phi = 35.1^\circ$, it corresponds to the orientation of the close-packed planes in the crystal lattice of the grain.

Periodic homogenization for generalized continua is used to predict the effective response of the polycrystal. The displacement field is assumed to be of the form

$$\mathbf{u}(x) = \mathbf{E} \cdot \mathbf{x} + \mathbf{v}(x), \quad (107)$$

with the fluctuation \mathbf{v} periodic, meaning that it takes identical values at homologous points of the unit cell (Forest et al. 2001). The plastic microdeformation field, χ^p , is assumed to be periodic, meaning that no rotational macroscopic plastic deformation is imposed to the unit cell. Its components are equal at homologous opposite nodes. According to periodic homogenization, the simple and double tractions \mathbf{t} and \mathbf{m} are antiperiodic at homologous points of the unit cell.

Polycrystals are random materials so that the periodicity constraint may lead to a bias in the estimation of the effective properties. This boundary effect can be alleviated by considering several realizations of the microstructure and performing ensemble averaging (Zeghadi et al. 2007).

Overall Cyclic Response of a Polycrystalline Aggregate

The finite element simulations of the boundary value problem presented previously have been conducted under generalized plane strain conditions on aggregates with a relatively small number of grains. The aim here is not to obtain a representative

Table 2 Sets of material parameters used in the 24-grain aggregate case (Fig. 11a). The intrinsic length scale, $l_s = \sqrt{A/H_\chi}$, is given for each set

Set	μ [MPa]	τ_c [MPa]	H_χ [MPa]	A [MPa mm ²]	l_s [μ m]
a	35,000	40	$3.0 \cdot 10^6$	$1.0 \cdot 10^{-2}$	$5.8 \cdot 10^{-2}$
b	35,000	40	$1.0 \cdot 10^6$	$1.0 \cdot 10^{-2}$	$1.0 \cdot 10^{-1}$
c	35,000	40	$3.5 \cdot 10^5$	$1.0 \cdot 10^{-2}$	$1.7 \cdot 10^{-1}$
d	35,000	40	$8.8 \cdot 10^4$	$1.0 \cdot 10^{-3}$	$1.1 \cdot 10^{-1}$

response but to catch the grain size effects and to explore qualitatively the impact of different sets of material parameters. In this section, a virtual material is considered with various intrinsic length scales. The macroscopic stress–strain curve is obtained by applying a cyclic simple shear loading controlled by the average strain component E_{12} on the aggregate of 52 grains with $d = 0.2 \mu\text{m}$ and the set of material parameters labeled (c) in Table 2. The mean stress component Σ_{12} is then computed:

$$\Sigma_{12} = \frac{1}{V} \int_V \sigma_{12} dV, \quad E_{12} = \frac{1}{V} \int_V \varepsilon_{12} dV, \quad (108)$$

where V denotes each polycrystal unit cell. The simulated response displays the kinematic hardening effect produced by the *microcurl* model. The stress–strain curves shown in Fig. 12 prove that this kinematic hardening is size dependent: it increases for smaller grains. Note that the observed overall kinematic hardening has two distinct sources: the intragranular back stress induced by plastic strain gradients and the intergranular internal stress that originates from the grain to grain plastic strain incompatibilities. The latter contribution is also predicted by classical crystal plasticity models.

Figure 13 presents the effect of the mean grain size, d , on the macroscopic flow stress at 1% plastic strain in the 52-grain aggregate in a log–log diagram for different intrinsic length scales, l_s , introduced through the sets of material parameters (labeled a, b, c, and d) given in Table 2. The considered loading conditions are still a simple shear test with periodic boundary conditions. The curves exhibit two plateaus for large ($d > 20 \mu\text{m}$) and small ($d < 0.1 \mu\text{m}$) mean grain sizes with a transition domain in between. This tanh shape indicates that when d is large compared to the intrinsic length scale, l_s , strain gradient effects are small, and the kinematic hardening arising from the *microcurl* model vanishes. The model saturates when d is of the order of l_s or smaller. The transition domain exhibits a strong size dependence, the polycrystalline aggregate becoming harder for decreasing grain sizes. The position of the transition zone, the maximum extra stress (the distance between the two plateaus), and the scaling law exponent, m , in the size-dependent domain are controlled by the material parameters used in the model. The two latter effects are controlled by the coupling modulus, H_χ ; they both increase for higher values of H_χ as shown in Fig. 13. The scaling exponent is defined as the slope in the log–log diagram in the inflection domain, reflecting the scaling law:

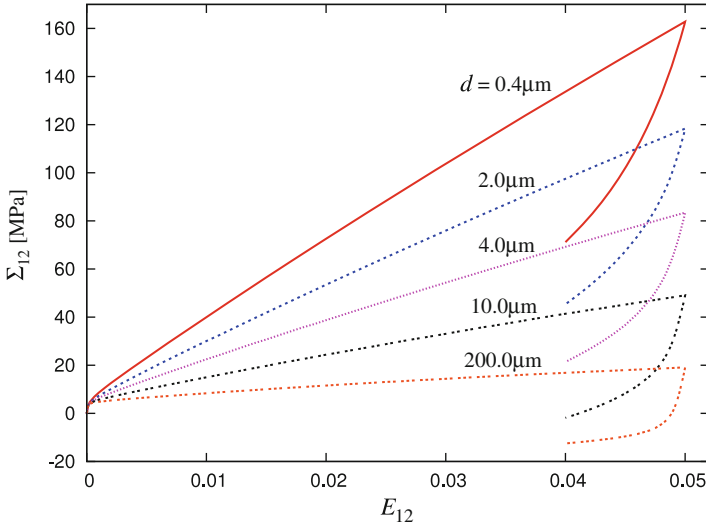


Fig. 12 Macroscopic stress–strain response of the 52-grain aggregate under simple shear for various mean grain sizes, d . The set of material parameters used is labeled (g) in Table 3

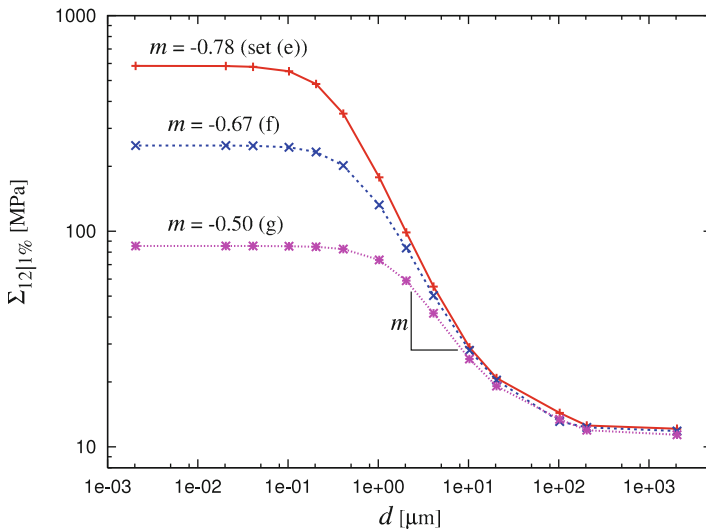


Fig. 13 Effect of the mean grain size, d , on the macroscopic flow stress, $\Sigma_{12}|_{1\%}$, at 1% plastic strain. The results are obtained for the 52-grain aggregate using the different sets of material parameters given in Table 3. The scaling law exponent, m , is identified in each case

$$\Sigma_{12} \propto d^m. \tag{109}$$

It is obtained with the sets of material parameters given in Table 2. The computed values range from -0.26 to -0.64 including the well-known Hall–Petch exponent

Table 3 Sets of material parameters used in the 52-grain aggregate case (Fig. 11b)

Set	μ [MPa]	τ_c [MPa]	Q [MPa]	b	$h^{\alpha\alpha}$	$h^{\alpha\beta, \alpha \neq \beta}$
e	27,000	0.75	7.9	10.2	1	4.4
f	27,000	0.75	7.9	10.2	1	4.4
g	27,000	0.75	7.9	10.2	1	4.4

Set	H_χ [MPa]	A [MPa mm ²]	l_s [μ m]
e	$1.0 \cdot 10^6$	$1.0 \cdot 10^{-2}$	$1.0 \cdot 10^{-1}$
f	$3.5 \cdot 10^5$	$1.0 \cdot 10^{-2}$	$1.7 \cdot 10^{-1}$
g	$5.0 \cdot 10^4$	$1.0 \cdot 10^{-2}$	$4.5 \cdot 10^{-1}$

$m = -0.5$. In fact, it was shown in Cordero et al. (2010a) that values of m ranging from 0 to -2 can be simulated with the *microcurl* model in the case of two-phase microstructures. In each case, these values are obtained without classical isotropic hardening, meaning that the linear kinematic hardening produced by the model is able to reproduce a wide range of scaling laws. Note that conventional strain gradient plasticity models do not lead to *tanh*-shape curves but rather to unbounded stress increase for vanishingly small microstructures (Cordero et al. 2010a).

Grain Size Effects in Idealized Aluminum Polycrystals

Similar finite element simulations have been performed on the idealized aluminum aggregate of 52 grains of Fig. 11. An additional isotropic hardening component is added as in (Méric et al. 1991) to obtain a more realistic response of large aluminum grains. The size-independent hardening law reads:

$$R^\alpha = \tau_c + Q \sum_{\beta}^n h^{\alpha\beta} (1 - \exp(-b\gamma_{\text{cum}}^\beta)), \quad (110)$$

where n is the number of slip systems (here $n = 2$), Q and b are material coefficients defining nonlinear isotropic hardening, $H^{\alpha\beta}$ is the interaction matrix, and $\gamma_{\text{cum}}^\beta$ is the accumulated micro-plastic slip on the slip system β . Cumulative plastic slip results from the integration of the differential equation $\dot{\gamma}_{\text{cum}}^\beta = |\dot{\gamma}^\beta|$. The material parameters used in these simulations are given in Table 3. The macroscopic stress–strain curves presented in Fig. 12 are obtained by applying a simple shear loading controlled by the average strain component E_{12} on the 52-grain aggregate with various mean grain sizes, d , taken in the size-dependent domain. The chosen set of material parameters has the label (g) in Table 3. These parameters are such that an acceptable description of aluminum polycrystals is obtained for large grains and that a Hall–Petch-like behavior is found in a plausible range of grain sizes. However, we did not attempt to calibrate the amplitude of the extra hardening so that simulation

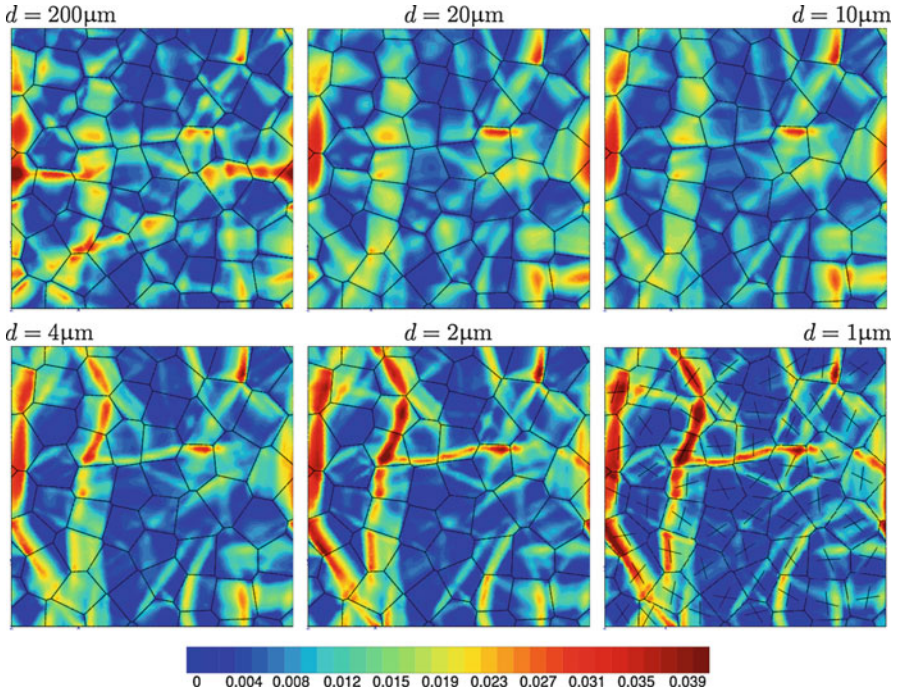


Fig. 14 Grain size effect on the accumulated plastic slip. These contour plots are obtained with the 52-grain aggregate for the same mean value of cumulative plastic strain $p = 0.01$. The set of material parameters (g) of Table 3 is used. The pairs of slip plane directions are represented for each grain on the $1 \mu\text{m}$ contour plot

predictions remain qualitative. The curves of Fig. 12 show again that the kinematic hardening produced by the model is strongly size dependent. The set of material parameters (g) of Table 3 gives the ideal Hall–Petch scaling law exponent $m = -0.5$.

An important output of the simulations is the dependence of the stress and strain fields in the grains of the polycrystal on grain size. Figures 14 and 15 show the contour plots of the field of accumulated plastic slip, computed as:

$$\dot{p} = \sqrt{\frac{2}{3} \dot{\boldsymbol{\varepsilon}}^p : \dot{\boldsymbol{\varepsilon}}^p}, \quad (111)$$

where $\boldsymbol{\varepsilon}^p$ is the symmetric part of the plastic deformation, \mathbf{H}^p , and the contour plots of the norm \mathbf{A} of the dislocation density tensor,

$$\|\mathbf{A}\| = \sqrt{\mathbf{A} : \mathbf{A}}, \quad (112)$$

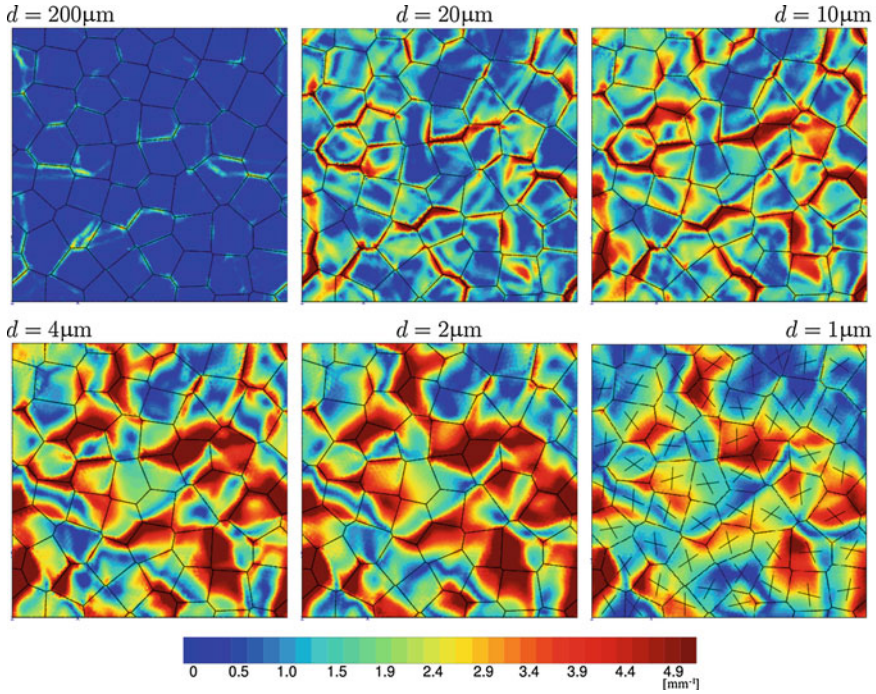


Fig. 15 Grain size effect on the norm of the dislocation density tensor. These contour plots are obtained with the 52-grain aggregate for the same mean value of macroscopic accumulated plastic strain $p = 0.01$. The set of material parameters (g) of Table 3 is used. The pairs of slip plane directions are represented for each grain on the 1 μm contour plot

respectively. The considered grain sizes are taken in the size-dependent domain where the evolution of the fields is assumed to be physically relevant. The chosen set of material parameters has the label (g) in Table 3; it corresponds to an intrinsic length scale $l_s = 0.45\mu\text{m}$ and gives a scaling law exponent $m = -0.5$. The mean value of the accumulated plastic slip is the same for all cases; only its distribution varies with the size of the microstructure as shown in Fig. 14.

The first contour plot of each figure is obtained for $d = 200\mu\text{m} \gg l_s = 0.45\mu\text{m}$, at the very beginning of the size-dependent behavior domain according to Fig. 13. At this size, the simulated fields show that p is quite inhomogeneous and that some deformation bands appear; $\|\mathbf{A}\|$ is localized at the grain boundaries and almost vanishes in the grain cores. The contour plots obtained for $2\mu\text{m} < d < 20\mu\text{m}$ show a significant evolution of both fields. One observes the formation of a network of strain localization bands with decreasing grain size. These bands are slip bands as they are parallel to the slip plane directions represented on the 1 μm contour plot of Fig. 14. They compensate the larger blue zones where plastic strain cannot develop due to the higher energy cost associated with its gradient. Plastic strain becomes stronger inside the localization bands. This is due to the fact that the contour plots

are given for fixed macroscopic cumulative plastic strain mean value of p , which implies that the applied total strain is higher for small grain sizes as suggested by Fig. 12. The field of the norm of the dislocation density tensor is still high close to grain boundaries and spreads over the grain cores. The last contour plot of each figure is obtained for $d = 1 \mu\text{m}$, a size close to l_s . Here the model starts to saturate, which can be seen from the simulated fields. The field of p does not evolve anymore and $\|\mathbf{A}\|$ decreases. In fact, as l_s controls the strain gradient effects, strong strain gradients cannot develop because they become energetically too expensive when the grain size is too small.

Conclusions

The micromorphic crystal plasticity theory introduces independent plastic microdeformation degrees of freedom. It represents a relaxation of the strain gradient plasticity model. It contains as a special case the micropolar crystal plasticity model. The advantage of the micromorphic framework is that it provides a wider range of modeling possibilities regarding constitutive laws and boundary conditions, as it was illustrated for the description of Hall–Petch effects in polycrystals. It has also merits regarding computational mechanical aspects since its implementation is rather straightforward and can be used for strain gradient plasticity computations based on proper internal constraints. The advantage of the micropolar theory compared to the micromorphic one is its reduced number of degrees of freedom, 3 instead of 8/9 in 3D (8 dof if plastic incompressibility is enforced). The micropolar model incorporates the effect of the lattice curvature tensor, which represents an essential part of the dislocation density tensor.

A full set of constitutive equations has been formulated for micromorphic crystal plasticity by extending the framework of generalized standard materials based on the introduction of free energy and dissipation potentials.

Simple examples of the plasticity of sheared single-crystal layers show the ability of the continuum models to reproduce the essential features of the collective behavior of dislocations piling up in thin layers, channels, and laminate microstructures, as predicted by discrete dislocation dynamics. The scaling laws predicted by the continuum models strongly depend on the choice of the constitutive equations. For example, standard quadratic free energy potential was shown to be inadequate to reproduce Orowan-like size effects. Strongly nonlinear potentials, including the logarithmic free energy density, were shown to be more closely related to the physics of dislocations.

Size effects also strongly influence the strain localization behavior of polycrystals for which the formation of intense slip bands is predicted in micron-size grains by the micromorphic models. This feature is to be related to the high energy cost of dislocation pileup formation in small grains.

Special attention was given to the cyclic response of crystals, which is of the utmost importance for the prediction of fatigue lifetime of materials. The question of recovery of plastic strain has been shown to be a central issue. In the absence

of strong forest interaction and cross-slip, total recovery of plastic strain gradients is possible, leading to nonconvex cyclic stress–strain loops as observed in some two-phase alloys. Note that results from discrete dislocation dynamics and the micropolar model were provided in the chapter dedicated to micropolar crystal plasticity. Similar comparisons can be found in the case of micromorphic model in the Reference Chang et al. (2016) where the continuum and discrete descriptions of dislocation pileups at interfaces are discussed.

Much work remains to be done in the development of predictive constitutive equations in micropolar and micromorphic crystal plasticity due to the complex underlying dislocation mechanisms. Especially, suitable modeling of grain boundary behavior remains a major issue for higher-order modeling of polycrystal plasticity.

Acknowledgments The first author is indebted to Dr. N.M. Cordero, Dr. S. Wulfinghoff, and Prof. E.B. Busso for their contribution to the presented micromorphic crystal plasticity theory. These contributions are duly cited in the references quoted in the text and listed below.

References

- E. Aifantis, The physics of plastic deformation. *Int. J. Plast.* **3**, 211–248 (1987)
- R.J. Asaro, Elastic–plastic memory and kinematic hardening. *Acta Metall.* **23**, 1255–1265 (1975)
- R. Asaro, Crystal plasticity. *J. Appl. Mech.* **50**, 921–934 (1983)
- Ashby, M., 1971. The deformation of plastically non-homogeneous alloys, in *Strengthening Methods in Crystals*, ed. by A. Kelly, R. Nicholson (Applied Science Publishers, London), pp. 137–192
- O. Aslan, N.M. Cordero, A. Gaubert, S. Forest, Micromorphic approach to single crystal plasticity and damage. *Int. J. Eng. Sci.* **49**, 1311–1325 (2011)
- V. Bennett, D. McDowell, Crack tip displacements of microstructurally small surface cracks in single phase ductile polycrystals. *Eng. Fract. Mech.* **70**(2), 185–207 (2003)
- V. Berdichevsky, On thermodynamics of crystal plasticity. *Scripta Mat.* **54**, 711–716 (2006a)
- V. Berdichevsky, On thermodynamics of crystal plasticity. *Scr. Mater.* **54**, 711–716 (2006b)
- P. Cermelli, M. Gurtin, On the characterization of geometrically necessary dislocations in finite plasticity. *J. Mech. Phys. Solids* **49**, 1539–1568 (2001)
- H.J. Chang, N.M. Cordero, C. Déprés, M. Fivel, S. Forest, Micromorphic crystal plasticity versus discrete dislocation dynamics analysis of multilayer pile-up hardening in a narrow channel. *Arch. Appl. Mech.* **86**, 21–38 (2016)
- W. Claus, A. Eringen, Three dislocation concepts and micromorphic mechanics, in *Developments in Mechanics. Proceedings of the 12th Midwestern Mechanics Conference*, vol. 6, (1969), pp. 349–358
- S. Conti, M. Ortiz, Dislocation microstructures and the effective behavior of single crystals. *Arch. Ration. Mech. Anal.* **176**, 103–147 (2005)
- N. Cordero, A. Gaubert, S. Forest, E. Busso, F. Gallerneau, S. Kruch, Size effects in generalised continuum crystal plasticity for two–phase laminates. *J. Mech. Phys. Solids* **58**, 1963–1994 (2010a)
- N.M. Cordero, A. Gaubert, S. Forest, E. Busso, F. Gallerneau, S. Kruch, Size effects in generalised continuum crystal plasticity for two-phase laminates. *J. Mech. Phys. Solids* **58**, 1963–1994 (2010b)
- N.M. Cordero, S. Forest, E. Busso, S. Berbenni, M. Cherkaoui, Grain size effects on plastic strain and dislocation density tensor fields in metal polycrystals. *Comput. Mater. Sci.* **52**, 7–13 (2012)

- L. De Luca, A. Garroni, M. Ponsiglione, Gamma-convergence analysis of Systems of Edge Dislocations: the self energy regime. *Arch. Ration. Mech. Anal.* **206**, 885–910 (2012)
- C. Déprés, C.F. Robertson, M.C. Fivel, Low-strain fatigue in aisi 316l steel surface grains: a three-dimensional discrete dislocation dynamics modelling of the early cycles i. Dislocation microstructures and mechanical behaviour. *Philos. Mag.* **84**(22), 2257–2275 (2004)
- Eringen, A., Claus, W., 1970. A micromorphic approach to dislocation theory and its relation to several existing theories, in *Fundamental Aspects of Dislocation Theory*, ed. by J. Simmons, R. de Wit, R. Bullough. National Bureau of Standards (US) Special Publication 317, vol. II (U.S. Government Printing Office, Washington, DC), pp. 1023–1062
- B. Fedelich, A microstructural model for the monotonic and the cyclic mechanical behavior of single crystals of superalloys at high temperatures. *Int. J. Mech. Sci.* **18**, 1–49 (2002)
- S. Forest, Some links between cosserat, strain gradient crystal plasticity and the statistical theory of dislocations. *Philos. Mag.* **88**, 3549–3563 (2008)
- S. Forest, The micromorphic approach for gradient elasticity, viscoplasticity and damage. *ASCE J. Eng. Mech.* **135**, 117–131 (2009)
- S. Forest, Nonlinear regularisation operators as derived from the micromorphic approach to gradient elasticity, viscoplasticity and damage. *Proc. R. Soc. A* **472**, 20150755 (2016)
- S. Forest, N. Guéinichault, Inspection of free energy functions in gradient crystal plasticity. *Acta Mech. Sinica.* **29**, 763–772 (2013) <https://doi.org/10.1007/s10409-013-0088-0>
- S. Forest, R. Sedláček, Plastic slip distribution in two-phase laminate microstructures: Dislocation-based vs. generalized-continuum approaches. *Philos. Mag. A* **83**, 245–276 (2003a)
- S. Forest, R. Sedláček, Plastic slip distribution in two-phase laminate microstructures: Dislocation-based vs. generalized-continuum approaches. *Philos. Mag. A* **83**, 245–276 (2003b)
- S. Forest, R. Sievert, Elastoviscoplastic constitutive frameworks for generalized continua. *Acta Mech.* **160**, 71–111 (2003)
- S. Forest, R. Sievert, Nonlinear microstrain theories. *Int. J. Solids Struct.* **43**, 7224–7245 (2006)
- S. Forest, F. Pradel, K. Sab, Asymptotic analysis of heterogeneous Cosserat media. *Int. J. Solids Struct.* **38**, 4585–4608 (2001)
- Forest, S., Ammar, K., Appolaire, B., Cordero, N., Gaubert, A., 2014. Micromorphic approach to crystal plasticity and phase transformation, in *Plasticity and Beyond*, ed. by J. Schroeder, K. Hackl. CISM International Centre for Mechanical Sciences, Courses and Lectures, no. 550 (Springer, Vienna), pp. 131–198
- M. Geers, R. Peerlings, M. Peletier, L. Scardia, Asymptotic behaviour of a pile-up of infinite walls of edge dislocations. *Arch. Ration. Mech. Anal.* **209**, 495–539 (2013)
- P. Germain, The method of virtual power in continuum mechanics. Part 2: microstructure. *SIAM J. Appl. Math.* **25**, 556–575 (1973)
- P. Grammenoudis, C. Tsakmakis, Micromorphic continuum part I: strain and stress tensors and their associated rates. *Int. J. Non-Linear Mech.* **44**, 943–956 (2009)
- I. Groma, F. Csikor, M. Zaiser, Spatial correlations and higher-order gradient terms in a continuum description of dislocation dynamics. *Acta Mater.* **51**, 1271–1281 (2003)
- I. Groma, G. Györgyi, B. Kocsis, Dynamics of coarse grain grained dislocation densities from an effective free energy. *Philos. Mag.* **87**, 1185–1199 (2007)
- M. Gurtin, A gradient theory of single-crystal viscoplasticity that accounts for geometrically necessary dislocations. *J. Mech. Phys. Solids* **50**, 5–32 (2002)
- M. Gurtin, L. Anand, Nanocrystalline grain boundaries that slip and separate: a gradient theory that accounts for grain-boundary stress and conditions at a triple-junction. *J. Mech. Phys. Solids* **56**, 184–199 (2008)
- M. Gurtin, L. Anand, Thermodynamics applied to gradient theories involving the accumulated plastic strain: the theories of Aifantis and Fleck and Hutchinson and their generalization. *J. Mech. Phys. Solids* **57**, 405–421 (2009)
- W. Han, B. Reddy, *Plasticity: Mathematical Theory and Numerical Analysis* (Springer, New York, 2013)
- C. Hirschberger, P. Steinmann, Classification of concepts in thermodynamically consistent generalized plasticity. *ASCE J. Eng. Mech.* **135**, 156–170 (2009)

- D.E. Hurtado, M. Ortiz, Surface effects and the size-dependent hardening and strengthening of nickel micropillars. *J. Mech. Phys. Solids* **60**(8), 1432–1446 (2012)
- D.E. Hurtado, M. Ortiz, Finite element analysis of geometrically necessary dislocations in crystal plasticity. *Int. J. Numer. Methods Eng.* **93**(1), 66–79 (2013)
- R. Kametani, K. Kodaera, D. Okumura, N. Ohno, Implicit iterative finite element scheme for a strain gradient crystal plasticity model based on self-energy of geometrically necessary dislocations. *Comput. Mater. Sci.* **53**(1), 53–59 (2012)
- Kröner, E., 1969. Initial studies of a plasticity theory based upon statistical mechanics, in *Inelastic Behaviour of Solids*, ed. by M. Kanninen, W. Adler, A. Rosenfield, R. Jaffee (McGraw-Hill, New York/London), pp. 137–147
- J. Lee, Y. Chen, Constitutive relations of micromorphic thermoplasticity. *Int. J. Eng. Sci.* **41**, 387–399 (2003)
- J. Mandel, Equations constitutives et directeurs dans les milieux plastiques et viscoplastiques. *Int. J. Solids Struct.* **9**, 725–740 (1973)
- L. Méric, P. Poubanne, G. Cailletaud, Single crystal modeling for structural calculations. Part 1: Model presentation. *J. Eng. Mat. Technol.* **113**, 162–170 (1991)
- S.D. Mesarovic, R. Baskaran, A. Panchenko, Thermodynamic coarsening of dislocation mechanics and the size-dependent continuum crystal plasticity. *J. Mech. Phys. Solids* **58**(3), 311–329 (2010)
- S. Mesarovic, S. Forest, J. Jaric, Size-dependent energy in crystal plasticity and continuum dislocation models. *Proc. R. Soc. A* **471**, 20140868 (2015)
- C. Miehe, S. Mauthe, F.E. Hildebrand, Variational gradient plasticity at finite strains. Part III: local-global updates and regularization techniques in multiplicative plasticity for single crystals. *Comput. Methods Appl. Mech. Eng.* **268**, 735–762 (2014)
- J. Nye, Some geometrical relations in dislocated crystals. *Acta Metall.* **1**, 153–162 (1953)
- N. Ohno, D. Okumura, Higher-order stress and grain size effects due to self-energy of geometrically necessary dislocations. *J. Mech. Phys. Solids* **55**, 1879–1898 (2007)
- N. Ohno, D. Okumura, Grain-size dependent yield behavior under loading, unloading and reverse loading. *Int. J. Mod. Phys. B* **22**, 5937–5942 (2008)
- M. Ortiz, E. Repetto, Nonconvex energy minimization and dislocation structures in ductile single crystals. *J. Mech. Phys. Solids* **47**(2), 397–462 (1999)
- H. Proudhon, W. Poole, X. Wang, Y. Bréchet, The role of internal stresses on the plastic deformation of the Al–Mg–Si–Cu alloy AA611. *Philos. Mag.* **88**, 621–640 (2008)
- B.D. Reddy, C. Wieners, B. Wohlmuth, Finite element analysis and algorithms for single-crystal strain-gradient plasticity. *Int. J. Numer. Methods Eng.* **90**(6), 784–804 (2012)
- R. Regueiro, On finite strain micromorphic elastoplasticity. *Int. J. Solids Struct.* **47**, 786–800 (2010)
- C. Sansour, S. Skatulla, H. Zbib, A formulation for the micromorphic continuum at finite inelastic strains. *Int. J. Solids Struct.* **47**, 1546–1554 (2010)
- P. Steinmann, Views on multiplicative elastoplasticity and the continuum theory of dislocations. *Int. J. Eng. Sci.* **34**, 1717–1735 (1996)
- R. Stoltz, R. Pelloux, Cyclic deformation and Bauschinger effect in Al–Cu–Mg alloys. *Scr. Metall.* **8**, 269–276 (1974)
- R. Stoltz, R. Pelloux, The Bauschinger effect in precipitation strengthened aluminum alloys. *Metallurgical. Transactions* **7A**, 1295–1306 (1976)
- B. Svendsen, S. Bargmann, On the continuum thermodynamic rate variational formulation of models for extended crystal plasticity at large deformation. *J. Mech. Phys. Solids* **58**(9), 1253–1271 (2010)
- R. Taillard, A. Pineau, Room temperature tensile properties of Fe-19wt.% Cr alloys precipitation hardened by the intermetallic compound NiAl. *Mater. Sci. Eng.* **56**, 219–231 (1982)
- S. Wulfinghoff, T. Böhlke, Equivalent plastic strain gradient enhancement of single crystal plasticity: theory and numerics. *Proc. R. Soc. A: Math. Phys. Eng. Sci.* **468**(2145), 2682–2703 (2012)

-
- S. Wulfinghoff, E. Bayerschen, T. Böhlke, A gradient plasticity grain boundary yield theory. *Int. J. Plast.* **51**, 33–46 (2013a)
 - S. Wulfinghoff, E. Bayerschen, T. Böhlke, Micromechanical simulation of the hall-petch effect with a crystal gradient theory including a grain boundary yield criterion. *PAMM* **13**, 15–18 (2013b)
 - S. Wulfinghoff, S. Forest, T. Böhlke, Strain gradient plasticity modeling of the cyclic behavior of laminate microstructures. *J. Mech. Phys. Solids* **79**, 1–20 (2015)
 - A. Zeghadi, S. Forest, A.-F. Gourgues, O. Bouaziz, Ensemble averaging stress–strain fields in polycrystalline aggregates with a constrained surface microstructure—part 2: crystal plasticity. *Philos. Mag.* **87**, 1425–1446 (2007)



Cosserat Approach to Localization in Geomaterials

19

Ioannis Stefanou, Jean Sulem, and Hadrien Rattetz

Contents

Introduction	688
The Cosserat Continuum	690
Linear and Angular Momentum Balance	691
Mass Balance	692
Energy Balance	693
Second Law of Thermodynamics	694
Phase Transitions Between the Species	694
Constitutive Laws for Cosserat Continua	695
Numerical Advantages of Cosserat Continuum	698
Mesh Dependency and Its Remedy	698
Shear Locking, Reduced Integration and Physical Hourglass Control	701
Upscaling and Homogenization	702
Rock Shear Layer of Cosserat Continuum: Fault Mechanics	704
Cauchy Continuum with Rate-Dependent Constitutive Law	705
Cosserat Elastoplasticity	706
Discussion: Rate-Dependent Models Versus Cosserat Continuum	707
Conclusion	707
References	708

Abstract

A renewed interest toward Cosserat or micropolar continuum has driven researchers to the development of specific models for upscaling discrete media such as masonry, granular assemblies, fault gouges, porous media, and biomaterials. Cosserat continuum is a special case of what is called micromorphic, generalized or higher-order continua. Due to the presence of

I. Stefanou · J. Sulem (✉) · H. Rattetz
Navier (CERMES), UMR 8205, Ecole des Ponts, IFSTTAR, CNRS, Champs-sur-Marne, France
e-mail: ioannis.stefanou@enpc.fr; jean.sulem@enpc.fr

internal lengths in its formulation, Cosserat continuum is quite attractive for addressing problems involving strain localization. It enables modeling the shear band thickness evolution, tracking the postlocalization regime, and correctly dissipating the energy when using numerical schemes. In this chapter, we summarize the fundamental governing equations of a Cosserat continuum under multiphysical couplings. Several examples of the numerical advantages of Cosserat continuum are also presented regarding softening behavior, strain localization, finite element formulation, reduced integration, and hourglass control. The classically used constitutive models in Cosserat elastoplasticity are presented and some common approaches for upscaling and homogenization in Cosserat continuum are discussed. Finally, a simple illustrative example of the adiabatic shearing of a rock layer under constant shear stress is presented in order to juxtapose a rate-independent Cosserat with a rate-dependent Cauchy formulation as far as it concerns strain localization.

Keywords

Cosserat continuum · Plasticity · Bifurcation theory · Multiphysical couplings · Strain localization · Shear locking · Upscaling · Faults

Introduction

Cosserat or micropolar continuum is a special case of what is called micromorphic, generalized, or higher-order continua. The general theory of micromorphic continua is general enough to describe various heterogeneous systems with microstructure. In Fig. 1, we briefly outline the various higher-order continuum theories, according to Germain (1973), and their special cases. Besides the classical continuum (also called Cauchy or Boltzmann continuum) and the Cosserat continuum, special cases of micromorphic continua are also the second gradient and the couple stress continuum theories.

In the last three decades, a renewed interest toward the Cosserat continuum has driven researchers to the development of specific continuous models for upscaling discrete media such as masonry, granular assemblies, fault gouges, porous media, and biomaterials (see, for instance, Besdo 1974; Anderson and Lakes 1994; Forest et al. 1999, 2001; Forest and Sab 1998; Stefanou et al. 2010; Sulem et al. 2011; Trovalusci et al. 2015; Masiani and Trovalusci 1996; Mühlhaus and Vardoulakis 1988; Bardet and Vardoulakis 2001; Pasternak and Mühlhaus 2006; Pasternak et al. 2006; de Buhan and Sudret 2000; among others). Cosserat continuum has seemed motivated by the advantages enclosed in its enhanced kinematics and the nonsymmetry of the stress tensor. When used for the formulation of continuum equivalent models for discrete media, Cosserat continuum allows to efficiently take into account high deformation gradients, relative particles' rotation, and scale effects. Moreover, Cosserat continuum enables the investigation of the phenomenon of wave dispersion, which characterizes the dynamic response of discrete media when the wavelength is comparable to the size of the microstructure (internal length).

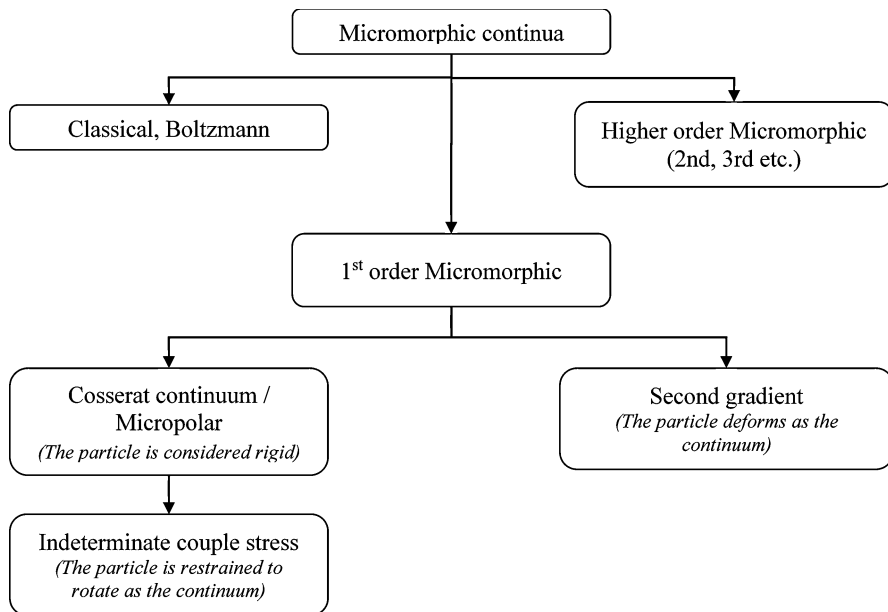


Fig. 1 Higher-order continuum theories according to Germain's terminology (Mindlin 1964; Eringen 1999)

Due to the presence of internal lengths in its formulation, Cosserat continuum is quite attractive for addressing problems involving strain localization. After the pioneering work of Mühlhaus and Vardoulakis (1988), who explained mathematically the shear band thickness and its evolution under shearing, Cosserat continuum became the first and nowadays one of the common modeling approaches for regularization of the ill-posed Cauchy continuum (Zervos et al. 2001; Papanastasiou and Zervos 2004; Papanastasiou and Vardoulakis 1992; De Borst and Sluys 1991; Godio et al. 2016a). Moreover, the correct simulation of the shear band thickness evolution (e.g., the thickness of the principal slip zone in a fault gouge) and of the postpeak behavior of a system, which Cosserat theory enables, guarantees correct energy dissipation.

This chapter is structured as follows. In section “[The Cosserat Continuum](#),” we give the general equations of a Cosserat continuum. In particular, the linear and angular momentum balance as well as the mass and energy balance equations and the second law of thermodynamics are extended for the Cosserat continuum. Under multiphysical couplings, involving chemistry and phase transformations, the above equations have to be adapted for momentum, mass, and energy exchange between the considered phases. Section “[The Cosserat Continuum](#)” is completed with a small paragraph discussing the additional terms due to this exchange. In section “[Constitutive Laws for Cosserat Continua](#),” we present some constitutive laws that are commonly used in Cosserat elasticity and elastoplasticity,

including softening and multiphysical couplings. In section “[Numerical Advantages of Cosserat Continuum](#),” we show the numerical advantages of Cosserat continuum theory for solving problems when softening behavior and strain localization are encountered. The commonly used approaches for upscaling and homogenization of heterogeneous structures in the frame of Cosserat continuum are discussed in section “[Upscaling and Homogenization](#).” Finally, in section “[Rock Shear Layer of Cosserat Continuum: Fault Mechanics](#),” a simple example of the adiabatic shearing of a rock layer under constant shear stress is presented in order to highlight the similarities of a rate-independent Cosserat and a rate-dependent Cauchy continuum as far it concerns the conditions for which shear band localization takes place. It is worth mentioning that Cosserat continuum is a promising framework in fault mechanics as it enables to compute the thickness of the principal slip zone inside a fault core, its dependence upon the microstructure of the fault gouge (Sulem et al. 2011; Rattetz et al. 2016a, b; Veveakis et al. 2013; Sulem and Stefanou 2016) and its impact on the seismic energy budget.

The Cosserat Continuum

Cosserat continuum equations can be derived in two different, but equivalent, ways, each one based on a different physical ansatz. The first one is based on the linear and angular momentum balance and the fundamental assumption of the existence of the so-called couple stresses at the material point (Fig. 2). The physical meaning

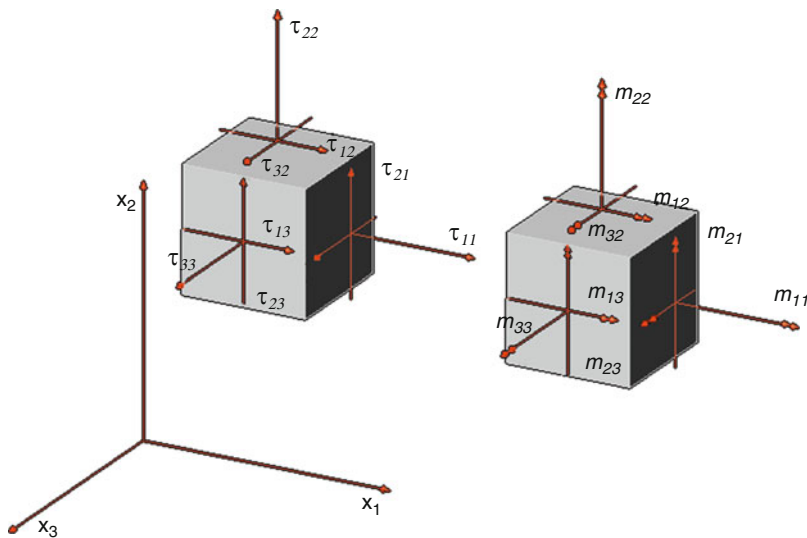


Fig. 2 Stresses and couple-stresses (moments) of Cosserat continuum (Stefanou et al. 2008)

of couple stresses, as well of their conjugate in energy kinematic measures, i.e., the gradient of the rotations (i.e., the curvatures), is a long-standing point of discussions in the scientific community and it is application-wise. For instance, in the Timoshenko theory of beams, which coincides with a one-dimensional Cosserat continuum, the Cosserat moment is simply the resultant section moment.

The second, alternative approach for deriving the Cosserat continuum equations is by postulating the form of the energy density function and by enriching directly the kinematics of the material point. Both ways are equivalent, but here we prefer to present the first one as only the momentum balance equations are needed. The energy of a Cosserat continuum can be then derived directly from the variational form of the Cosserat equilibrium equations.

Linear and Angular Momentum Balance

Let us consider a solid of volume V and its boundary ∂V . The linear momentum is given as follows:

$$P_i = \int_V \rho v_i dV \quad (1)$$

where v_i is the velocity of a material point, $i = 1,2,3$ (in three dimensions), and ρ is the density. The resultant force is written as:

$$F_i = \int_V f_i dV + \int_{\partial V} t_i dS \quad (2)$$

where f_i is a body force per unit volume and t_i is the traction vector, with $t_i = \tau_{ij}n_j$ and n_j the normal unit vector at the boundary. Repeated indices denote summation over the repeated index. Applying the Newton's law $\frac{DP_i}{Dt} = F_i$, we obtain:

$$\tau_{ij,j} + f_i = \rho \frac{Dv_i}{Dt} \quad (3)$$

In the case of a multiphase material, the velocity of each constituent phase has to be considered separately. Moreover, the momentum equations have to be written separately for each phase and then combined to give again Eq. 3, but with $\rho \frac{Dv_i}{Dt} = (1-n)\rho^s \frac{D^s v_i^s}{Dt} + n\rho^f \frac{D^f v_i^f}{Dt}$, where, here, f stands for the fluid phase and s for the solid one, n is the Eulerian porosity (pore volume per unit volume of porous material in actual state), and $\frac{D^x(\cdot)}{Dt} = \frac{\partial(\cdot)}{\partial t} + (\cdot)_{,j} v_j^x$. The above equation is valid as far as no exchange or transformation takes place between the phases.

The moment of momentum about a point is:

$$H_i = \int_V e_{ijk} x_j \rho v_k dV + \int_V I_{ij} \omega_j dV \tag{4}$$

where e_{ijk} is the Levi-Civita perturbation symbol, I_{ij} the microinertia tensor, and ω_j the Cosserat rotational velocity of the material point. In the special case of an isotropic microinertia tensor, $I_{ij} = I_s \delta_{ij}$, where δ_{ij} is the Kronecker delta.

The resultant moment about the same point is written as:

$$L_i = \int_V (e_{ijk} x_j f_k + \mu_i) dV + \int_S e_{ijk} x_j t_k dS + \int_V m_i dS \tag{5}$$

where μ_i is a body couple stress per unit volume and $m_i = m_{ij} n_j$ is the couple stress tractor, with m_{ij} the couple stress tensor. Euler's law $\frac{DH_i}{Dt} = L_i$, yields:

$$m_{ij,j} + e_{ijk} \tau_{kj} = I_{ij} \frac{D\omega_j}{Dt} \tag{6}$$

From the above equation, we observe that the stress tensor τ_{ij} is not symmetric due to the presence of the couple stress gradient and the microinertia terms. τ_{ij} can be decomposed into a symmetric σ_{ij} and an antisymmetric part s_{ij} : $\tau_{ij} = \sigma_{ij} + s_{ij}$. In this case, the symmetric part plays the role of the Cauchy stress tensor.

Mass Balance

The derivation of the pore pressure diffusion-generation equation from the mass balance follows the classical derivation procedure as in the classical Cauchy continuum. In particular, it becomes:

$$\frac{\partial p^f}{\partial t} = \Lambda \frac{\partial T}{\partial t} - \frac{1}{\beta^*} \frac{\partial \varepsilon}{\partial t} + \frac{1}{\beta^*} w_{i,i} \tag{7}$$

where $w_i = n (v_i^f - v_i^s)$ is the Darcy's fluid (relative) velocity, $\frac{\partial \varepsilon}{\partial t} = v_{i,i}$ is the volumetric strain rate, $\Lambda = \frac{\lambda^*}{\beta^*}$ is the undrained pore fluid pressurization coefficient, with $\lambda^* = n\lambda^f + (1-n)\lambda^s$, where λ^f is the fluid thermal expansion coefficient and λ^s is the solid thermal expansion coefficient, $\beta^* = n\beta^f + (1-n)\beta^s$ is the storage coefficient, where β^f is the fluid compressibility and β^s is the compressibility of the solid phase.

Assuming Darcy's law for the fluid flow, the above equation becomes:

$$\frac{\partial p^f}{\partial t} = \Lambda \frac{\partial T}{\partial t} - \frac{1}{\beta^*} \frac{\partial \varepsilon}{\partial t} + c_{hy} P_{,ii}^f \tag{8}$$

where $c_{hy} = \frac{k}{\eta_f \beta^*}$ is the hydraulic diffusivity with k the hydraulic permeability, which is assumed here isotropic and constant and, η_f the fluid viscosity.

Energy Balance

Similar to the mass balance equation, the derivation of the energy balance equation and in a second step of the heat equation follows the same steps as in the case of the Cauchy continuum. However, the work of the external, generalized forces and the kinetic energy have some additional terms due to the Cosserat rotations and moments. More specifically, assuming a micropolar solid phase and a nonmicropolar fluid, the kinetic energy rate is:

$$\begin{aligned} \frac{DK}{Dt} &= \frac{DK^s}{Dt} + \frac{DK^f}{Dt} \\ &= \frac{D}{Dt} \left\{ \int_V \left[\frac{1}{2} (1-n) \rho^s v_i^s v_i^s + \frac{1}{2} (1-n) I^s \omega_i \omega_i \right] dV \right. \\ &\quad \left. + \int_V \frac{1}{2} n \rho^f v_i^f v_i^f dV \right\} \end{aligned} \quad (9)$$

The internal energy is written as follows:

$$\begin{aligned} \frac{DE}{Dt} &= \frac{DE^s}{Dt} + \frac{DE^f}{Dt} \\ &= \frac{D}{Dt} \int_V \left[(1-n) \rho^s \mathcal{E}^s + n \rho^f \mathcal{E}^f \right] dV \end{aligned} \quad (10)$$

here \mathcal{E}^s and \mathcal{E}^f are the internal energy of the solid and fluid phase, respectively, per unit mass.

The work rate of external forces for a Cosserat continuum is given by:

$$\frac{DW}{Dt} = \int_{\partial V} \left(t_i^s v_i^s + m_i \omega_i + t_i^f v_i^f \right) dS + \int_V \left(f_i^s v_i^s + \mu_i \omega_i + f_i^f v_i^f \right) dV \quad (11)$$

Notice that due to the presence of two phases in the medium, the volumetric as well as the external generalized forces and tractions are split over the two phases. For a chemically inert system and in absence of any exchange between the solid and the fluid phase, the equations remain the same with Eqs. 3 and 6. According to the first law of thermodynamics, we have:

$$\frac{DK}{Dt} + \frac{DE}{Dt} = \frac{DQ}{Dt} + \frac{DW}{Dt} \quad (12)$$

where $\frac{DQ}{Dt} = - \int_{\partial V} h_i n_i dS = - \int_V h_{i,i} dV$ is the rate of heat input (given to V through its boundary ∂V) and h_i is the heat flux vector components. After some algebraic

manipulations, the energy balance equation becomes:

$$\begin{aligned} & (1-n)\rho^s \frac{D^s \mathcal{E}^s}{Dt} + n\rho^f \frac{D^f \mathcal{E}^f}{Dt} + \rho^f \mathcal{E}_{,i}^f w_i \\ & - \tau_{ij} \dot{\gamma}_{ij} - m_{ij} \dot{\kappa}_{ij} + h_{i,i} \\ & + (p^f w_i)_{,i} - \rho^f \left(f_i - \frac{D^f v_i^f}{Dt} \right) w_i = 0 \end{aligned} \quad (13)$$

where $\dot{\gamma}_{ij} = v_{i,j} + e_{ijk} \omega_k$ and $\dot{\kappa}_{ij} = \omega_{i,j}$. Neglecting the Cosserat effects and adapting the notation, the above equation coincides with Houlsby and Puzrin (2000) and Coussy (2004).

Assuming an elastoplastic constitutive law (see section “[Constitutive Laws for Cosserat Continua](#)”), the energy balance equation results to the heat equation for Cosserat continua:

$$\frac{\partial T}{\partial t} = c_{th} \frac{\partial^2 T}{\partial x_i \partial x_i} + \frac{1}{\rho C} \tau_{ij} \dot{\gamma}_{ij}^p + \frac{1}{\rho C} m_{ij} \dot{\kappa}_{ij}^p \quad (14)$$

where T is the temperature, $\dot{\gamma}_{ij}^p$ the plastic deformation rate tensor, and $\dot{\kappa}_{ij}^p$ the plastic curvature rate tensor and c_{th} the thermal diffusivity coefficient. For deriving the above equation, small deformations and an isotropic Fourier’s law were considered.

Second Law of Thermodynamics

Following the same derivations, the second law of thermodynamics yields:

$$\int_V \left[\frac{D^s S}{Dt} - (S - S^f) (\rho^f w_i)_{,i} + S_{,i}^f \rho^f w_i + \left(\frac{h_i}{T} \right)_{,i} \right] dV \geq 0 \quad (15)$$

where S denotes the specific entropy of the mixture per unit mass and S^x the specific entropy of the specie x . The above equation coincides with the one of (Houlsby and Puzrin 2000) after proper substitutions and algebra. Under specific assumptions for the constitutive behavior of the solid and the fluid phases, the above equation can be further developed.

Phase Transitions Between the Species

Phase transitions from the solid to the fluid phase and vice versa, due to, for instance, dissolution/precipitation mechanisms or dehydration, has as a result the addition of supplementary terms in the above equations. In particular, if $\dot{r}_{s \rightarrow f} = -\dot{r}_{f \rightarrow s}$ expresses the rate of transition from the solid to the fluid phase, Eqs. 3, 6, 8, 13, and 15 become:

$$\tau_{ij,j} + f_i + \left(v_i^f - v_i^s\right) \dot{r}_{s \rightarrow f} = \rho \frac{Dv_i}{Dt} \quad (16)$$

$$m_{ij,j} + e_{ijk} \tau_{kj} + \alpha \omega_i \dot{r}_{s \rightarrow f} = I_{ij} \frac{D\omega_j}{Dt} \quad (17)$$

$$\frac{\partial p^f}{\partial t} = \Lambda \frac{\partial T}{\partial t} - \frac{1}{\beta^*} \frac{\partial \varepsilon}{\partial t} + \frac{1}{\beta^*} w_{i,i} - c_{\text{ch}} \dot{r}_{s \rightarrow f} \quad (18)$$

$$\begin{aligned} (1-n) \rho^s \frac{D^s \varepsilon^s}{Dt} + n \rho^f \frac{D^f \varepsilon^f}{Dt} + \rho^f \mathcal{E}_{,i}^f w_i \\ - \tau_{ij} \dot{\gamma}_{ij} - m_{ij} \dot{\kappa}_{ij} + h_{i,i} \\ + (p^f w_i)_{,i} - \rho^f \left(f_i - \frac{D^f v_i^f}{Dt} \right) w_i \\ + \dot{r}_{s \rightarrow f} \frac{1}{2} \left[\left(v_i^f - v_i^s \right) \left(v_i^f - v_i^s \right) + \alpha \omega_i \omega_i + 2 \left(\varepsilon^f - \varepsilon^s \right) \right] = 0 \end{aligned} \quad (19)$$

$$\begin{aligned} \int_V \left[\frac{D^s S}{Dt} - (S - S^f) (\rho^f w_i)_{,i} + S_{,i}^f \rho^f w_i + \left(\frac{h_i}{T} \right)_{,i} \right] dV \\ + \int_V \dot{r}_{s \rightarrow f} (S^f - S^s) dV \geq 0 \end{aligned} \quad (20)$$

where $c_{\text{ch}} = \frac{1}{\beta^*} \frac{\rho^f - \rho^s}{\rho^f \rho^s}$ is the chemical pressurization coefficient and $\alpha = \frac{I_s}{\rho^s}$. The chemical pressurization coefficient may be insignificant in most of the cases of dissolution/precipitation reactions in rocks (Sulem and Stefanou 2016). However, this is not the case for dehydration or decarbonation reactions (Sulem and Famin 2009; Brantut et al. 2011).

Constitutive Laws for Cosserat Continua

Deriving constitutive relations for Cosserat continuum is not a trivial task. This is owed to the fact that Cosserat theory has intrinsic internal lengths and consequently the constitutive law is not scale invariant. In the simplest case, i.e., under small deformations, in a centrosymmetric, linear, elastic, isotropic Cosserat medium, the stresses are related to the generalized elastic deformation measures according to the following constitutive relations (Vardoulakis 2009):

$$\begin{aligned} \tau_{ij} &= K \gamma_{kk}^{\text{el}} \delta_{ij} + 2G \left(\gamma_{(ij)}^{\text{el}} - \frac{1}{3} \gamma_{kk}^{\text{el}} \delta_{ij} \right) + 2\eta_1 G \gamma_{[ij]}^{\text{el}} \\ m_{ij} &= 4GR^2 \left(\kappa_{(ij)}^{\text{el}} + \eta_2 \delta_{ij} \kappa_{kk}^{\text{el}} \right) + 4GR^2 \eta_3 \kappa_{[ij]}^{\text{el}} \end{aligned} \quad (21)$$

where K is the bulk modulus, G is the shear modulus, η_1, η_2, η_3 are positive material constants, and R is an internal length parameter, which for granular materials is often identified to the mean radius of the grains in the representative volume element (RVE). $\gamma_{(ij)}$ and $\gamma_{[ij]}$ denote, respectively, the symmetric and antisymmetric parts of $\gamma_{ij} = u_{i,j} + e_{ijk} \theta_k$, where u_i are the components of

Table 1 Coefficients used for the generalized J2 invariants for Cosserat continuum (Sulem and Vardoulakis 1990)

	2D model	3D model
Static consideration	$h_i = \{\frac{3}{4}, \frac{-1}{4}, 1, 0\} g_i = \{\frac{3}{2}, \frac{1}{2}, 1, 0\}$	$h_i = \{\frac{2}{3}, \frac{-1}{6}, \frac{2}{3}, \frac{-1}{6}\} g_i = \{\frac{8}{5}, \frac{2}{5}, \frac{8}{5}, \frac{2}{5}\}$
Kinematical consideration	$h_i = \{\frac{3}{8}, \frac{1}{8}, \frac{1}{4}, 0\} g_i = \{3, -1, 4, 0\}$	$h_i = \{\frac{2}{5}, \frac{1}{10}, \frac{2}{5}, \frac{1}{10}\} g_i = \{\frac{8}{3}, \frac{-2}{3}, \frac{8}{3}, \frac{-2}{3}\}$

the displacement vector and θ_k the components of the Cosserat rotation vector (infinitesimal rotations). The Cosserat shear modulus, which expresses the stiffness related to the relative rotation of the particle (e.g. of a grain) with respect to the macrorotation of the continuum (e.g., assemblage of grains) is defined as $G_c = \eta_l G$.

In a small deformations framework, the classical J2 plasticity criteria can be generalized for Cosserat continua. This approach is based on the generalization of the stress and strain invariants in order to account for the Cosserat couple stresses, curvatures, and rotations. Based on micromechanical considerations and averaging, the deviatoric stress and strain invariants take the following form (Rattez et al. 2016b; Sulem and Vardoulakis 1990; Vardoulakis and Sulem 1995):

$$\begin{aligned}
 q &= \sqrt{h_1 \tau_{ij}^d \tau_{ij}^d + h_2 \tau_{ij}^d \tau_{ji}^d + \frac{1}{R^2} (h_3 m_{ij}^d m_{ij}^d + h_4 m_{ij}^d m_{ji}^d)} \\
 \gamma &= \sqrt{g_1 \gamma_{ij}^d \gamma_{ij}^d + g_2 \gamma_{ij}^d \gamma_{ji}^d + R^2 (g_3 \kappa_{ij}^d \kappa_{ij}^d + g_3 \kappa_{ij}^d \kappa_{ji}^d)}
 \end{aligned}
 \tag{22}$$

where $\tau_{ij}^d, m_{ij}^d, \gamma_{ij}^d,$ and κ_{ij}^d are the deviatoric parts of the stress, couple-stress, strain, and curvature tensors respectively. $h_p,$ and g_p ($p = 1, \dots, 4$) are coefficients that can be determined by micromechanical considerations on the kinematics or the statics of an assemblage of grains in contact (Sulem and Vardoulakis 1990; Mühlhaus and Vardoulakis 1987) (see Table 1).

Notice that the internal length R appears directly in the above expressions. Once the above invariants are derived, various plasticity criteria can be used. We refer, for example, to the von Mises, the Drucker-Prager, or to the Cam-clay criterion, which are commonly used in metal, soil, and rock mechanics and they are expressed in terms of the q, p invariants. Volumetric and shear hardening parameters can be also defined allowing not only to reproduce the stress-strain response of a geomaterial, but also the shear band thickness evolution (Mühlhaus and Vardoulakis 1987), its thickness, and the postbifurcation regime, which is crucial in many applications as it is related to stress redistribution, collapse, and energy dissipation. In a small strains framework, the generalized deformation rate measures are decomposed in elastic and plastic parts as follows:

$$\begin{aligned}
 \dot{\gamma}_{ij} &= \dot{\gamma}_{ij}^{el} + \dot{\gamma}_{ij}^{pl} \\
 \dot{\kappa}_{ij} &= \dot{\kappa}_{ij}^{el} + \dot{\kappa}_{ij}^{pl}
 \end{aligned}
 \tag{23}$$

Assuming a Drucker-Prager yield surface and plastic potential:

$$F = q + \mu p \leq 0 \quad (24)$$

$$Q = q + \beta p \quad (25)$$

where $p = \tau_{kk}$, μ is the friction coefficient and β the dilatancy coefficient; the elastoplastic incremental generalized stress-strain relationships become:

$$\begin{aligned} \dot{\tau}_{ij} &= C_{ijkl}^{\text{ep}} \dot{\gamma}_{kl} + D_{ijkl}^{\text{ep}} \dot{\kappa}_{kl} + E_{ijkl}^{\text{ep}} \dot{T} \delta_{kl} \\ \dot{\mu}_{ij} &= M_{ijkl}^{\text{ep}} \dot{\kappa}_{kl} + L_{ijkl}^{\text{ep}} \dot{\gamma}_{kl} + N_{ijkl}^{\text{ep}} \dot{T} \delta_{kl} \end{aligned} \quad (26)$$

where

$$\begin{aligned} C_{ijkl}^{\text{ep}} &= C_{ijkl}^e - \frac{\langle 1 \rangle}{H_p} b_{ij}^Q b_{kl}^F & D_{ijkl}^{\text{ep}} &= -\frac{\langle 1 \rangle}{H_p} b_{ij}^Q b_{kl}^{\text{FM}} \\ E_{ijkl}^{\text{ep}} &= -\left(C_{ijkl}^e - \frac{\langle 1 \rangle}{H_p} b_{ij}^Q b_{kl}^F \right) & L_{ijkl}^{\text{ep}} &= -\frac{\langle 1 \rangle}{H_p} b_{ij}^Q b_{kl}^F \\ M_{ijkl}^{\text{ep}} &= M_{ijkl}^e - \frac{\langle 1 \rangle}{H_p} b_{ij}^Q b_{kl}^{\text{FM}} & N_{ijkl}^{\text{ep}} &= \frac{\langle 1 \rangle}{H_p} b_{ij}^Q b_{kl}^F \end{aligned} \quad (27)$$

C_{ijkl}^e and M_{ijkl}^e are the elastic constitutive tensors for the stress and the couple stress, respectively, derived from Eq. 21, and

$$\begin{aligned} b_{kl}^F &= \frac{\partial F}{\partial \tau_{ij}} C_{ijkl}^e; & b_{ij}^Q &= C_{ijkl}^e \frac{\partial F}{\partial \tau_{kl}}; & b_{kl}^{\text{FM}} &= \frac{\partial F}{\partial \mu_{ij}} M_{ijkl}^e \\ H_p &= \frac{\partial F}{\partial \tau_{ij}} C_{ijkl}^e \frac{\partial Q}{\partial \tau_{kl}} + \frac{\partial F}{\partial \mu_{ij}} M_{ijkl}^e \frac{\partial Q}{\partial \mu_{kl}} + H_s \end{aligned} \quad (28)$$

H_s is the hardening modulus due to shear plastic deformations, defined by:

$$\frac{\partial F}{\partial \gamma^p} = -H_s \quad (29)$$

Once the constitutive parameters of the material are calibrated, a bifurcation analysis can provide important information regarding strain localization evolution under shearing (Mühlhaus and Vardoulakis 1987). Various multiphysical phenomena (Sulem et al. 2011; Veveakis et al. 2012, 2013) as well as grain cataclasis (Rattez et al. 2016b) can also be considered. Predicting strain localization is of paramount importance in geotechnical applications involving strongly non-associated materials (common case) and in fault mechanics where the multiphysical couplings, energy dissipation, softening behavior, and grain size evolution play the central role for earthquake triggering and coseismic slip (see next sections for more details).

It is worth emphasizing that the internal length(s) that are present in the Cosserat continuum theory have somehow to be determined or identified. This can be done

either by using bottom-up, upscaling, and homogenization techniques that consider the microstructure in detail (see section “Upscaling and Homogenization”) or through classical geotechnical tests. However, in practice of classical geotechnical testing, the calibration of these parameters is ignored.

Numerical Advantages of Cosserat Continuum

Mesh Dependency and Its Remedy

It is well known that when it comes to strain localization, the governing equations based on a classical Boltzmann continuum become ill-posed and when numerical approaches are used for integration of these equations, the result is mesh dependent (e.g., De Borst et al. 1993). With the term mesh dependency, we mean that the numerical results depend on the size and the type of the finite elements used. More specifically, the global behavior depends on the choice of the mesh and by refining it no convergence of the numerical results is observed. In Fig. 3 we

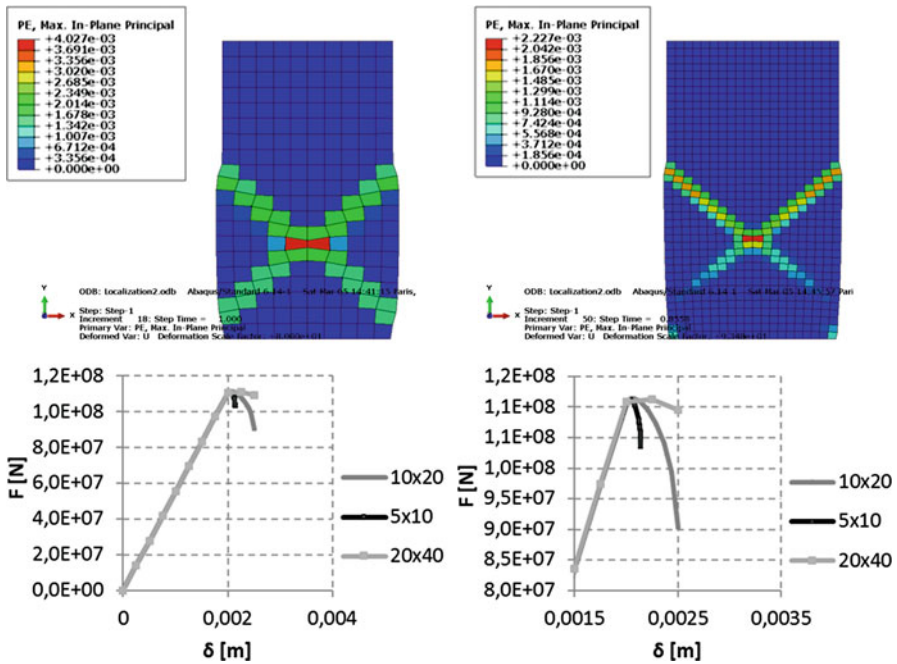


Fig. 3 Biaxial test of a softening material for different mesh sizes. The developed plastic deformations, the shear band thickness, and the total applied force – displacement depends on the mesh size

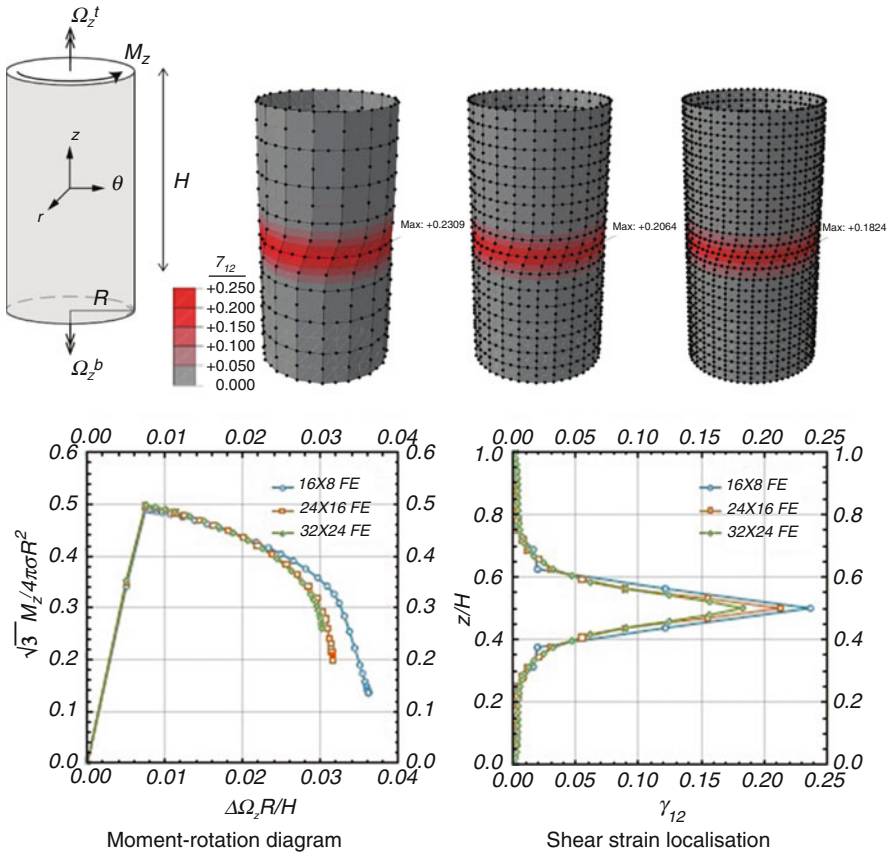


Fig. 4 Above: Deformation pattern and isocontour of in-plane shear strain distribution for different discretizations of the hollow cylinder. Below: Macroscopic behavior of the hollow cylinder (left) and localization of the in-plane shear strain (right) for different discretizations with the COSS8R finite element (FE) (Godio et al. 2015)

present this mathematical and numerical artifact, which is owed to the absence of internal lengths in the classical continuum. Moreover, the developed shear bands (contracting or dilating) are always one to two finite elements thick.

Cosserat continuum is one of the available techniques for regularizing mesh dependency during shear banding (De Borst and Sluys 1991; De Borst et al. 1993; De Borst 1984). In Fig. 4, we present the simulations that were carried out using COSS8R finite element, a homemade User Element for Abaqus, which is currently extended for use to any classical finite element code (Godio et al. 2016a). The problem treated is a thin hollow cylinder made of a softening material under shearing. Three different finite element discretizations were tested. In the same

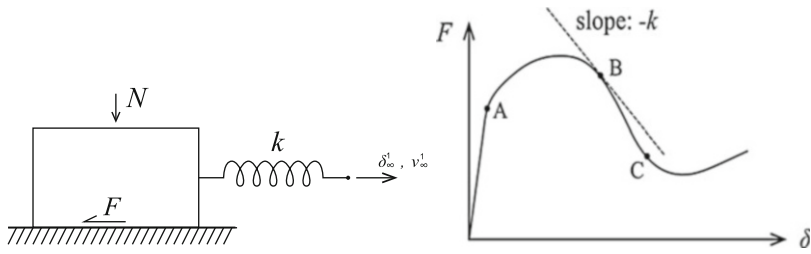


Fig. 5 *Left:* The classical spring-slider model used for understanding earthquake nucleation and energy partition (Scholz 2002; Kanamori and Brodsky 2004). *Right:* Schematic representation of the mechanical behavior of a rock interface (fault). From A to B, the slip is aseismic. Instability takes place at B (coseismic slip). The role of the softening branch is apparent

figure, we present the resulting macroscopic behavior of the cylinder in terms of the normalized reaction moment and the differential axial rotation. The onset of plastic deformations rapidly culminates with a softening branch, which is accompanied by strain localization.

Strain localization occurs exactly as in the shear layers studied by De Borst (1991). The curves show important mesh-independency, i.e., the thickness (width) of the localization zone does not depend on the number of elements falling inside this zone and the total force-axial rotation curve is independent of the mesh size. This is not only a fundamental feature of finite elements based on the Cosserat continuum, but it demonstrates also the fast convergence of the chosen Finite Element interpolation scheme (Godio et al. 2015).

Besides failure analysis in geotechnical engineering (Goodman 1989), the post-peak behavior of the geomaterial plays a crucial role in earthquake nucleation. In the idealistic, basic case of considering a fault zone as an interface with no multiphysical couplings and assuming a rate-independent behavior for the rock material, i.e., $F = F(\delta)$, where F is the interface friction and δ the slip (Fig. 5 left), bifurcation theory leads (see Mam and Quennehen 2016; Stefanou and Alevizos 2016) to the classical relationship for earthquake nucleation: $dF(\delta)/d\delta \leq -k$ (Scholz 2002). This condition is presented graphically in Fig. 5 (right). Instabilities are triggered (or induced) not at the peak, but in the softening branch, i.e., at B–C. Before that, there is slip but it is *aseismic*. This shows the importance of correctly capturing the postpeak behavior of the rock material (numerically and analytically). Moreover, correct simulation of the shear band thickness evolution (e.g., the thickness of the principal slip zone in a fault) guarantees that the dissipation of energy through the various multiphysical mechanisms that take place (e.g., temperature and pore pressure increase, shear heating, thermal pressurization, thermally induced chemical reactions, grain size evolution, etc.) is correctly calculated before and during the seismic slip.

Shear Locking, Reduced Integration and Physical Hourglass Control

In the above paragraph, we described the advantages of Cosserat continuum for inelastic behavior and strain localization compared to the classical continuum. Yet, Cosserat continuum numerical advantages do not stop there.

Shear locking is a common numerical problem when lower-order interpolation functions are used in finite element analysis. Shear locking is the phenomenon of extremely slow convergence of the numerical solution to the exact one, even when extremely fine discretizations are used. Due to the chosen interpolation functions, a large ratio of the energy input is erroneously transformed to shear strain energy instead of flexural (Zienkiewicz et al. 1971) resulting to a stiffer response. A numerical trick that it is commonly used in finite elements is reduced integration of the stiffness matrix. This alleviates shear locking without losing accuracy (in an incremental formulation, the right-hand side is completely/correctly integrated). However, this numerical trick comes with a price. Reduced integration of common Cauchy continuum finite elements (e.g., linear or quadratic polynomial interpolations) leads to a deficiency of the stiffness matrix. In other words, the stiffness matrix has some eigenvalues that are zero and consequently some deformation modes are associated to zero strain energy (spurious modes). For linear and quadratic Cauchy finite elements, these zero-energy modes have the geometrical shape of an hourglass. In traditional finite element technology, a small artificial/numerical stiffness is added to these deformation modes (represented by the corresponding to the zero eigenvalues eigenvectors). This technique is called hourglass control. However, when important stress gradients take place, hourglass control can fail.

Cosserat continuum offers a natural way to avoid hourglass control. It can be shown (Godio et al. 2015) that after reduced integration of the stiffness matrix, no zero eigenvalues appear in the stiffness matrix. In Fig. 6, we consider the well-discussed plane foundation problem (Zienkiewicz et al. 2013). Due to reduced integration, an hourglass mode is activated in the rigid element on top as a result of the applied force. This hourglass mode is able to propagate within a certain area

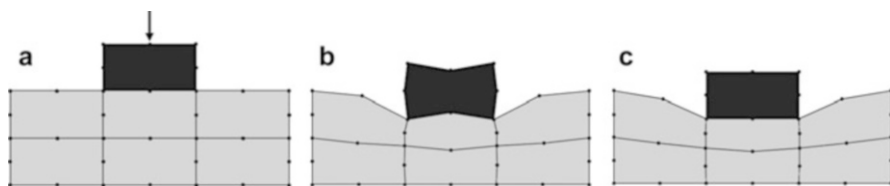


Fig. 6 Foundation of a rigid footing: (a) a highly rigid element resting upon a layer of flexible elements with reduced integration; (b) propagating zero-energy mode in an assembly of Cauchy (S8R) elements; (c) accurate results from a COSS8R subdivision (Godio et al. 2015)

in the elements below. On the contrary, the COSS8R element does not have this deficiency, as the rotational DOF perform an intrinsic action of hourglass control and no artificial hourglass control is needed.

Upscaling and Homogenization

Homogenization or upscaling methods is a class of methods that aim at deriving an equivalent continuum theory that describes the macroscopic behavior of heterogeneous systems, i.e., of systems with microstructure. (Asymptotic) homogenization is a mathematically rigorous, well-established theory for performing this task (Sánchez-Palencia 1986; Bakhvalov and Panasenko 1989; Pinho-da-Cruz et al. 2009; Charalambakis 2010). This method is based on the asymptotic expansion of the various state fields (displacements, deformations, stresses) in terms of a small quantity ε , which represents the ratio of the characteristic size of the ‘elementary volume’ over the overall size of the structure, and provides an equivalent to the heterogeneous system homogeneous continuum for $\varepsilon \rightarrow 0$. Besides the rigorous mathematical formulation of this approach, its main advantage is the ability to determine error estimators of the derived continuum for finite values of ε . However, when it comes to generalized continua, such as the Cosserat continuum, that possess internal lengths, the asymptotic limit $\varepsilon \rightarrow 0$ loses interest as it cancels out these internal length (Pradel and Sab 1998). In other words, by this pass to the limit, asymptotic homogenization erases any internal length that are related to the material’s microstructure, which higher-order continuum theories, such as Cosserat, are in principle able to capture.

To overcome this problem, several alternative schemes have been proposed in the literature for upscaling heterogeneous systems (see Anderson and Lakes 1994; Forest and Sab 1998; Stefanou et al. 2008, 2010; Trovalusci et al. 2015; Bardet and Vardoulakis 2001; Pradel and Sab 1998; Sab and Pradel 2009; Rezakhani and Cusatis 2016; Godio et al. 2016b; Bacigalupo and Gambarotta 2011, 2012; Baraldi et al. 2016, among others). The majority of these schemes is based upon the homogeneous equivalent continuum concept (see Charalambakis 2010), in the sense that the derived Cosserat continuum shares (a) the same energy (internal energy, dissipation) and (b) the same kinematics with the heterogeneous, discrete medium. The classical asymptotic homogenization expansion ansatz that leads to a Cauchy continuum as the ratio of the size of the unit cell over the overall structure tends to zero is not followed in these approaches. Therefore, these heuristic approaches remain applicable even when the size of the microstructure is not infinitesimal as compared to the overall size of the system, or, in other words, when scale separation is no more applicable.

A typical example for applying and testing these upscaling methods is masonry-like structures. Masonry can be seen as a geomaterial whose building blocks are often quasi-periodically arranged in space. Moreover, the building blocks are at the human scale, which makes them an ideal toy model, contrary to granular media whose microstructure is small, shows geometric complexity and has to be statis-

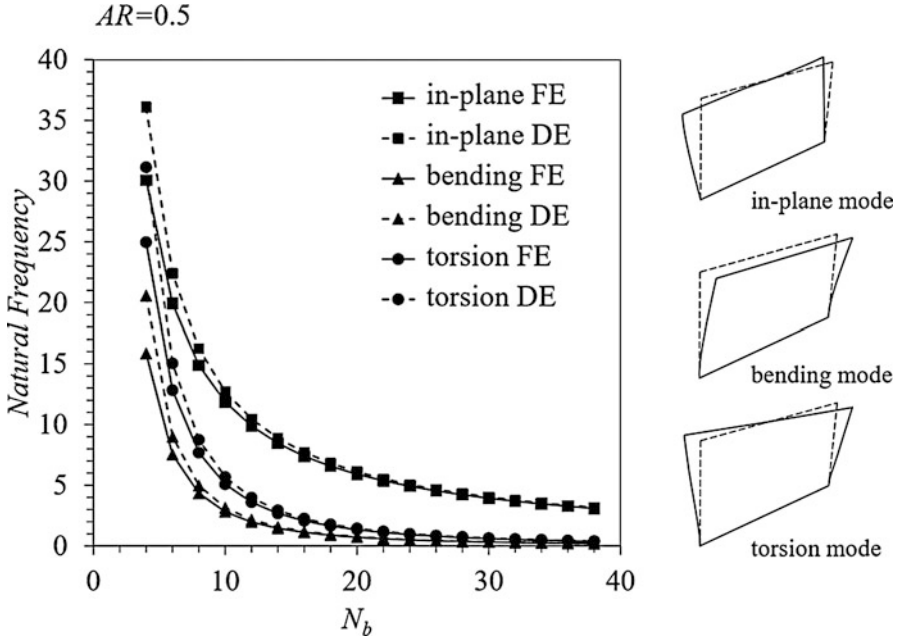


Fig. 7 Modal frequencies of a masonry panel versus the number of building blocks: comparison between the results extracted by discrete element method and by the use of the Cosserat homogenized plate model for masonry (Godio et al. 2016a; Stefanou et al. 2008)

tically described (Stefanou and Sulem 2014, 2016). When the upscaling scheme is correctly formulated, it is possible to capture the wave dispersion behavior of a heterogeneous system even when the wave length is comparable to the block size. Notice that in this case, the classical homogeneous Cauchy continuum approach fails as it is not a dispersive medium. In Fig. 7, we present the modal frequencies of a masonry panel which was up-scaled with Cosserat continuum (Stefanou et al. 2008; Godio et al. 2015) in function of the number of its building blocks. Even when the number of the building blocks is small, the Cosserat homogenized continuum model succeeds in representing the dynamics of the discrete heterogeneous structure. In Fig. 8, we present the out-of-plane-displacement contours of the first three flexural modal shapes of a homogenized masonry panel and a comparison with the flexural modal shapes provided by discrete elements (Godio et al. 2014).

In the case of granular media, the approach described in Bardet and Vardoulakis (2001, 2003) or in Rezakhani and Cusatis (2016) can be followed for upscaling and determining the effective parameters of an equivalent Cosserat continuum. It is worth pointing out that these upscaling techniques may provide valuable information on strain localization and energy dissipation in the absence of detailed experimental data as it is the case of the complex behavior of fault gouges. However, Cosserat continuum is effective only under shearing. In the case of pronounced extension or compaction, the Cosserat additional terms have no effect and at least a first-order micromorphic continuum has to be used (Fig. 1).

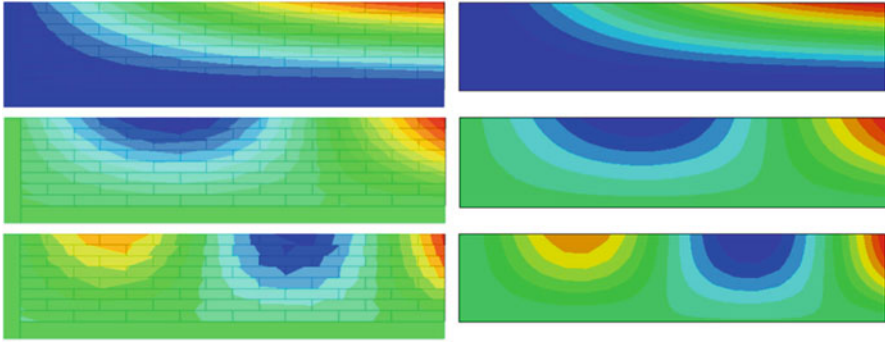


Fig. 8 Out-of-plane-displacement contours of the first three flexural modal shapes. *Left* Discrete elements solution. *Right* Homogenized Cosserat finite element solution (Godio et al. 2014)

Rock Shear Layer of Cosserat Continuum: Fault Mechanics

In this section, we give a simple example of Cosserat theory used for modeling the adiabatic shearing of a rock layer under constant shear stress. Two different modeling frameworks of at least, at first approximation, different physical assumptions are juxtaposed and compared as far as it concerns the conditions for which shear band localization takes place.

The first framework is the Cauchy continuum with rate-dependent constitutive law (viscoplasticity). Rate-dependent constitutive models and, in general, viscoplastic constitutive laws in the frame of Cauchy continuum are frequently used as, under some conditions, they can lead to finite thickness shear band formation. The relation between viscosity and shear band thickness (and consequently material length scale) has been discussed in several publications (e.g., Wang et al. 1996). The second modeling framework is Cosserat elastoplasticity.

Thermal softening is taken into account as a destabilizing mechanism that may lead to shear band localization. The complexity of the chosen constitutive laws and of the multiphysical couplings considered is kept to a minimum degree in order to reveal the salient features of each framework and highlight their similarities or their differences regarding strain localization. For a more detailed modeling in the frame of Cosserat continuum involving thermo-poro-chemo-mechanical couplings and more elaborate constitutive laws for the rock material, the reader is referred to Sulem et al. (2011), Godio et al. (2016a), and Veveakis et al. (2012, 2013). The reader is referred to (Rice et al. 2014) for the Cauchy rate-dependent framework under thermo-poro-chemo-mechanical couplings.

The thickness of the rock layer is D and constant normal and shear stresses are applied at its boundaries as depicted in Fig. 9. Initially, the layer is considered to be in a state of homogeneous shear deformation.

In both models, it is assumed that all the plastic work is converted to heat and that Fourier's law is applicable. Under these assumptions, the heat equation is written in indicial notation as follows (see Eq. 14):

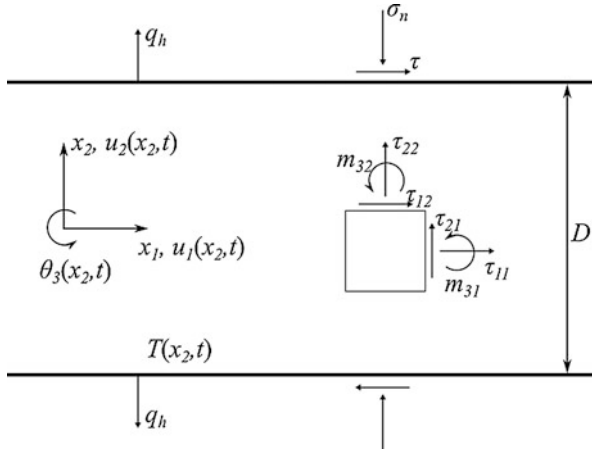


Fig. 9 Shearing of a rock layer: Cosserat rotational degree of freedom ω and couple stresses

$$\frac{\partial T}{\partial t} = c_{th} \frac{\partial^2 T}{\partial x_2^2} + \frac{1}{\rho C} \tau_{ij} \dot{\gamma}_{ij}^p + \frac{1}{\rho C} m_{ij} \dot{\kappa}_{ij}^p \tag{30}$$

where the infinite layer hypothesis was considered and consequently the derivatives in the x_1 and x_3 directions vanish (invariance). Small deformations are considered and the slip event is sufficiently rapid in order to justify adiabatic conditions at the boundaries of the layer. In the case of the Cauchy continuum, the third term in the right-hand side vanishes and the stress tensor is symmetric.

Cauchy Continuum with Rate-Dependent Constitutive Law

Let's assume a simple rate-dependent constitutive law for the shear stress at a point inside the shear layer:

$$\sigma_{12} = \tau_0 + H \dot{\gamma}_{12} + \xi (T - T_s) \tag{31}$$

where H is a mechanical hardening parameter (positive), ξ a thermal softening parameter (negative), T_s a reference temperature, and τ_0 the shear stress at steady state and reference temperature.

For a Cauchy continuum, the linear momentum balance is:

$$\sigma_{ij,j} = 0 \tag{32}$$

Inertia terms and body forces are neglected in this example. The angular momentum balance imposes the symmetry of the stress tensor, $\sigma_{ij} = \sigma_{ji}$. For a Cauchy continuum, $\gamma_{ij} = u_{i,j}$, where u_i is the displacement in the i th direction.

At steady state $T = T^* = T_s, \sigma_{12} = \sigma_{12}^* = \tau_0, \sigma_{22} = \sigma_{22}^* = \sigma_0, \dot{\gamma}_{12} = \dot{\gamma}_{12}^* = 0,$ and $\dot{T}^* = 0.$ This state will be stable as long as any perturbation does not grow in time. By perturbing the temperature and displacements fields at steady state, $T = T^* + \tilde{T}, u_i = u_i^* + \tilde{u}_i$ and by neglecting higher order terms, Eqs. 30, 31, and 32 become:

$$\begin{aligned} \tilde{\sigma}_{12} &= H \dot{\gamma}_{12}^p + \xi \tilde{T}; & \frac{\partial \tilde{\sigma}_{12}}{\partial x_2} &= 0; & \frac{\partial \tilde{\sigma}_{22}}{\partial x_2} &= 0 \\ \frac{\partial \tilde{T}}{\partial t} &= c_{th} \frac{\partial^2 \tilde{T}}{\partial x_2^2} + \frac{1}{\rho C} \sigma_{12}^* \dot{\gamma}_{12}^p \end{aligned} \tag{33}$$

The perturbations $\tilde{T}, \tilde{u}_i,$ should fulfill the boundary conditions of the rock layer. Equation 33 together with the boundary conditions $\frac{\partial \tilde{T}}{\partial z} \Big|_{z=\pm \frac{D}{2}} = 0, \tilde{\sigma}_{12} (z = \pm \frac{D}{2}) = 0$ and $\tilde{\sigma}_{22} (z = \pm \frac{D}{2}) = 0$ form a linear system of partial differential equations, which admits solutions of the type $\tilde{u}_i = U_i e^{st} \sin \frac{2\pi}{\lambda} z, \tilde{T} = Te^{st} \cos \frac{2\pi}{\lambda} z,$ where s is the so-called growth coefficient and $\lambda = \frac{D}{N}, N = 1, 2, 3, \dots$. By replacing into Eq. 33 we obtain: $s = -\frac{\xi \tau_0}{H \rho C} - \frac{4\pi^2 c_{th}}{\lambda^2}$. The system is unstable when $s > 0$ or, equivalently, when the wavelength of the perturbation is bigger than a critical wavelength $\lambda_{cr}^{rd} : \lambda > \lambda_{cr}^{rd} = 2\pi \sqrt{\frac{H c_{th} \rho C}{-\xi \tau_0}}.$

Cosserat Elastoplasticity

An elastic perfectly plastic constitutive behavior with thermal softening is assumed in this example. More advanced Cosserat constitutive models such as the M\u00fchlhaus-Vardoulakis Cosserat plasticity model (Vardoulakis and Sulem 1995) might be used (see section “Constitutive Laws for Cosserat Continua”), but the advantage of this simple model is that analytical derivations are simple to perform, which permits a convenient comparison with the above rate-dependent model. By analogy with the Cauchy rate-dependent model presented in the previous paragraph, the yield surface is defined as:

$$F = \tau_{(ij)} - \tau_0 - \xi (T - T_s) \leq 0 \tag{34}$$

where $\tau_{(ij)}$ denotes the symmetric part of the stress tensor. In this way, the same shear stress limit and thermal softening with the Cauchy model is retrieved if one neglects the rate-dependent term in Eq. 31. Nevertheless, because of the chosen yield surface (Eq. 34), the plastic curvatures are zero and therefore they do not contribute to the heat equation (Eq. 30). A centrosymmetric, linear elastic isotropic Cosserat medium is considered (see Eq. 21).

At steady state, we have a Cauchy continuum under homogeneous shear. In particular, $\dot{T}^* = 0, T = T^* = T_s, \tau_{(12)} = \tau_{(12)}^* = \tau_0, \tau_{[12]} = \tau_{[12]}^* = 0, m_{32} = m_{32}^* = 0,$ and $\sigma_{22} = \sigma_{22}^* = \sigma_0.$ This state will be stable as long as any perturbation does not grow in time. The temperature, the displacement, and the rotation fields

at steady state are perturbed ($T = T^* + \tilde{T}$, $u_i = u_i^* + \tilde{u}_i$, $\theta_3 = \theta_3^* + \tilde{\theta}_3$) as in the Cauchy case. The perturbations \tilde{T} , \tilde{u}_i , and $\tilde{\theta}_3$ have to fulfill the boundary conditions of the rock layer as in the Cauchy continuum case and additionally $\tilde{m}_{32}(z = \pm \frac{D}{2}) = 0$. A linear system is then formed which admits solutions of the form: $\tilde{u}_i = U_i e^{st} \sin \frac{2\pi}{\lambda} z$, $\tilde{\theta}_3 = \Theta_3 e^{st} \cos \frac{2\pi}{\lambda} z$, $\tilde{T} = T e^{st} \cos \frac{2\pi}{\lambda} z$. The critical growth coefficient is then: $s = -\frac{16G\pi^4 R^2 \rho C c_{th}}{\lambda^2 (4G\pi^2 R^2 \rho C + (8\pi^2 R^2 + \lambda^2) \xi \tau_0)}$ where we set $G_c = G$ for simplicity. The system is unstable when $s > 0$ or, equivalently when the wavelength of the perturbation is bigger than a critical wavelength λ_{cr}^{Cos} : $\lambda > \lambda_{cr}^{Cos} = 2\pi \sqrt{\frac{R^2(G\rho C + 2\xi\tau_0)}{-\xi\tau_0}} \approx 2\pi \sqrt{\frac{R^2 G \rho C}{-\xi\tau_0}}$. For typical values of the shear modulus, the applied shear stress at the boundary, the thermal softening parameter, and specific heat, it holds $G\rho C \gg \xi\tau_0$.

Discussion: Rate-Dependent Models Versus Cosserat Continuum

Even though both frameworks are based on different constitutive assumptions and micromechanisms, the resemblance of the expressions for the critical wavelength impels an analogy between the hardening parameter of the viscoplastic model H and the Cosserat internal length, which here is chosen equal to the mean grain radius: $Hc_{th} \sim R^2 G$. We observe that in the rate-dependent framework, a characteristic length appears from the combination of strain rate hardening and thermal diffusivity, whereas in the Cosserat framework, it is directly related to the material grain size. The hardening parameter H can be measured experimentally for a given rock and it generally decreases during shearing together with the size of the grains and the shear modulus, which also decrease due to important shearing and comminution. It is worth mentioning that the term $R^2 G$ represents the rolling stiffness of the grains, which, in comparison with the classical Cauchy continuum, rigidifies the system in the same way that the viscous term in the rate-dependent friction law does. If we take the example of a highly granulated fault gouge with a grain size of $10 \mu\text{m}$ and assuming a shear modulus $G = 300 \text{ MPa}$, then for $c_{th} = 1 \text{ mm}^2/\text{s}$, the hardening parameter H is equal to $H = 0.03 \text{ MPa s}$, which is in agreement with experimental measurements (Blanpied et al. 1995; Chester and Higgs 1992). Consistently, we observe the similar role of the diffusion length and of the Cosserat internal length in the control of the thickness of the localization zone.

Conclusion

In this chapter, we summarized the fundamental, balance equations of a Cosserat continuum under multiphysical couplings that involve a fluid and a solid phase. In the beginning, the fluid was considered not to react with the solid phase in order to present the basic thermo-hydro-mechanical framework. The additional terms due

to fluid momentum, mass, and energy exchange were presented next in the case of chemically reactive fluids (thermo-hydro-chemo-mechanical couplings) resulting to phase transitions. Under appropriate, application-wise constitutive laws, the above equations can be used for modeling various physical systems by taking into account the size of the microstructure in a continuum mechanics framework.

Cosserat continuum has several advantages. First of all, it remedies the mesh dependency of the classical Cauchy continuum and leads to correct energy dissipation in inelasticity and softening. Notice that the latter is of particular importance in rock mechanics as it controls stress redistribution and failure, in fault mechanics as it determines the seismic energy budget, and in soil mechanics especially when nonassociative plastic flow behavior is involved. Several examples are presented in section “[Numerical Advantages of Cosserat Continuum.](#)” Moreover, Cosserat continuum assures in a physical manner hourglass control in reduced integration finite element formulations.

Finally, a simple example of the adiabatic shearing of a rock layer under constant shear stress is presented in order to highlight the similarities of a rate-independent Cosserat and a rate-dependent Cauchy continuum as far it concerns strain localization conditions. Cosserat continuum is a promising framework for studying many problems where microstructure is of paramount importance. Faults is one of them (Sulem et al. 2011; Rattetz et al. 2016a, b; Veveakis et al. 2013; Sulem and Stefanou 2016; Brantut et al. 2017).

References

- W.B. Anderson, R.S. Lakes, Size effects due to Cosserat elasticity and surface damage in closed-cell polymethacrylimide foam. *J. Mater. Sci.* **29**, 3–9 (1994)
- A. Bacigalupo, L. Gambarotta, Computational two-scale homogenization of periodic masonry: characteristic lengths and dispersive waves. *Comput. Methods Appl. Mech. Eng.* **213–216**, 16–28 (2012)
- A. Bacigalupo, L. Gambarotta, Multi-scale modelling of periodic masonry: characteristic lengths and dispersive waves, in *AIMETA 2011: 20th Conference of the Italian Association of Theoretical and Applied Mechanics*. University of Bologna. ISBN 978-88-906340-0-0 (2011)
- N. Bakhvalov, G. Panasenko, *Homogenisation: Averaging Processes in Periodic Media: Mathematical Problems in the Mechanics of Composite Materials* (Springer, Berlin, 1989)
- D. Baraldi, S. Bullo, A. Cecchi, Continuous and discrete strategies for the modal analysis of regular masonry. *Int. J. Solids Struct.* **84**, 82–98 (2016)
- J.P. Bardet, I. Vardoulakis, The asymmetry of stress in granular media. *Int. J. Solids Struct.* **38**(2), 353–367 (2001)
- J.P. Bardet, I. Vardoulakis, Reply to discussion by Dr. Katalin Bagi. *Int. J. Solids Struct.* **40**, 7683–7683 (2003)
- D. Besdo, Ein Beitrag zur nichtlinearen Theorie des Cosserat-Kontinuums. *Acta Mech.* **20**, 105–131 (1974)
- M.L. Blanpied, D.A. Lockner, J.D. Byerlee, Frictional slip of granite at hydrothermal conditions. *J. Geophys. Res.* **100**(B7), 13045 (1995)
- N. Brantut, J. Sulem, A. Schubnel, Effect of dehydration reactions on earthquake nucleation: stable sliding, slow transients, and unstable slip. *J. Geophys. Res. Solid Earth* **116**, 1–16 (2011) July 2010

- N. Brantut, I. Stefanou, J. Sulem, Dehydration-induced instabilities at intermediate depths in subduction zones. *J. Geophys. Res. Solid Earth* **122**(8), 6087–6107 (2017)
- P. de Buhan, B. Sudret, Micropolar multiphase model for materials reinforced by linear inclusions. *Eur. J. Mech. A/Solids* **19**(4), 669–687 (2000)
- N. Charalambakis, Homogenization techniques and micromechanics. A survey and perspectives. *Appl. Mech. Rev.* **63**(3), 30803 (2010)
- F.M. Chester, N.G. Higgs, Multimechanism friction constitutive model for ultrafine quartz gouge at hypocentral conditions. *J. Geophys. Res.* **97**(B2), 1859 (1992)
- O. Coussy, *Poromechanics* (Wiley, West Sussex, 2004)
- R. de Borst, Possibilities and limitations of finite elements for limit analysis. *Géotechnique* (2), 199–210 (1984)
- R. de Borst, Simulation of strain localization: a reappraisal of the Cosserat continuum. *Eng. Comput.* **8**(4), 317–332 (1991)
- R. de Borst, L. J. Sluys, Localisation in a Cosserat continuum under static and dynamic loading conditions. *Comput. Methods Appl. Mech. Eng.* **90**(1–3), 805–827 (1991)
- R. de Borst, L.J. Sluys, H.B. Mühlhaus, J. Pamin, Fundamental issues in finite element analyses of localization of deformation. *Eng. Comput.* **10**(2), 99–121 (1993)
- A.-C. Eringen, *Microcontinuum Field Theories* (Springer Verlag, New York, 1999)
- S. Forest, K. Sab, Cosserat overall modeling of heterogeneous materials. *Mech. Res. Commun.* **25**(4), 449–454 (1998)
- S. Forest, R. Dendievel, G.R. Canova, Estimating the overall properties of heterogeneous Cosserat materials. *Model. Simul. Mater. Sci. Eng.* **7**(5), 829–840 (1999)
- S. Forest, F. Pradel, K. Sab, Asymptotic analysis of heterogeneous Cosserat media. *Int. J. Solids Struct.* **38**(26–27), 4585–4608 (2001)
- P. Germain, The method of virtual power in continuum mechanics. Part 2: microstructure. *SIAM J. Appl. Math.* **25**(3), 556–575 (1973)
- M. Godio, I. Stefanou, K. Sab, J. Sulem, Cosserat elastoplastic finite elements for masonry structures. *Key Eng. Mater.* **624**, 131–138 (2014)
- M. Godio, I. Stefanou, K. Sab, J. Sulem, Dynamic finite element formulation for Cosserat elastic plates. *Int. J. Numer. Methods Eng.* **101**(13), 992–1018 (2015)
- M. Godio, I. Stefanou, K. Sab, J. Sulem, Multisurface plasticity for Cosserat materials: plate element implementation and validation. *Int. J. Numer. Methods Eng.* **108**(5), 456–484 (2016a)
- M. Godio, I. Stefanou, K. Sab, J. Sulem, S. Sakji, A limit analysis approach based on Cosserat continuum for the evaluation of the in-plane strength of discrete media: application to masonry. *Eur. J. Mech. A/Solids* **66**, 168–192 (2016b)
- R.E. Goodman, *Introduction to Rock Mechanics* (Wiley, New York, 1989), pp. 293–334
- G.T. Houlisby, A.M. Puzrin, A thermomechanical framework for constitutive models for rate-independent dissipative materials. *Int. J. Plast.* **16**(9), 1017–1047 (2000)
- H. Kanamori, E.E. Brodsky, The physics of earthquakes. *Rep. Prog. Phys.* **67**(8), 1429–1496 (2004)
- K. Mam S. Quennehen, Réactivation de faille par géothermie profonde Implémentation d'une loi de frottement sous ABAQUS. Master Thesis, Ecole des Ponts ParisTech (2016)
- R. Masiani, P. Trovalusci, Cosserat and Cauchy materials as continuum models of brick masonry. *Meccanica* **31**(4), 421–432 (1996)
- R.D. Mindlin, Micro-structure in linear elasticity. *Arch. Ration. Mech. Anal.* **16**(1) (1964)
- H.-B. Mühlhaus, I. Vardoulakis, The thickness of shear bands in granular materials. *Géotechnique* **37**(3), 271–283 (1987)
- H.B. Mühlhaus, I. Vardoulakis, The thickness of shear hands in granular materials. *Géotechnique* **38**(3), 331–331 (1988)
- P.C. Papanastasiou, I. Vardoulakis, Numerical treatment of progressive localization in relation to borehole stability. *Int. J. Numer. Anal. Methods Geomech.* **16**(6), 389–424 (1992)
- P. Papanastasiou, A. Zervos, Wellbore stability analysis: from linear elasticity to postbifurcation modeling. *Int. J. Geomech.* **4**(1), 2–12 (2004)
- E. Pasternak, A.V. Dyskin, H.B. Mühlhaus, Cracks of higher modes in Cosserat continua. *Int. J. Fract.* **140**(1), 189–199 (2006a)

- E. Pasternak, A.V. Dyskin, Y. Estrin, Deformations in transform faults with rotating crustal blocks. *Pure Appl. Geophys.* **163**(9), 2011–2030 (2006b)
- J. Pinho-da-Cruz, J.A. Oliveira, F. Teixeira-Dias, Asymptotic homogenisation in linear elasticity. Part I: mathematical formulation and finite element modelling. *Comput. Mater. Sci.* **45**(4), 1073–1080 (2009)
- F. Pradel, K. Sab, Cosserat modelling of elastic periodic lattice structures. *C. R. Acad. Sci. Ser. IIB Mech.* **326**(11), 699–704 (1998)
- H. Rattetz, I. Stefanou, J. Sulem, E. Veveakis, T. Poulet, Insight into the different couplings involved in the weakening of faults and their effect on earthquake nucleation, in *American Geophysical Union Fall Meeting* (2016a)
- H. Rattetz, I. Stefanou, J. Sulem, Thermo-hydro mechanical couplings and strain localisation in 3D continua with microstructure. Part I : theory and linear stability analysis, under Prep. (2016b)
- R. Rezakhani, G. Cusatis, Asymptotic expansion homogenization of discrete fine-scale models with rotational degrees of freedom for the simulation of quasi-brittle materials. *J. Mech. Phys. Solids* **88**, 320–345 (2016)
- J.R. Rice, J.W. Rudnicki, J.D. Platt, Stability and localization of rapid shear in fluid-saturated fault gouge: 1. Linearized stability analysis. *J. Geophys. Res. Solid Earth* **119**(5), 4311–4333 (2014)
- K. Sab, F. Pradel, Homogenisation of periodic Cosserat media. *Int. J. Comput. Appl. Technol.* **34**(1), 60 (2009)
- E. Sánchez-Palencia, Homogenization in mechanics. A survey of solved and open problems. *Rend. Sem. Mat. Univ* **44**, 1–46 (1986)
- C.H. Scholz, *The Mechanics of Earthquakes and Faulting*, 2nd edn. (Cambridge University Press, Cambridge, 2002)
- I. Stefanou S. Alevizos, Fundamentals of bifurcation theory and stability analysis, in *Modelling of Instabilities and Bifurcation in Geomechanics, ALERT Geomaterials Doctoral School 2016*, ed. by J. Sulem, I. Stefanou, E. Papamichos, E. Veveakis (Aussois, 2016), http://alertgeomaterials.eu/data/school/2016/2016_ALERT_schoolbook.pdf
- I. Stefanou, J. Sulem, Chemically induced compaction bands: triggering conditions and band thickness. *J. Geophys. Res. Solid Earth* **119**(2), 880–899 (2014)
- I. Stefanou, J. Sulem, Existence of a threshold for brittle grains crushing strength: two-versus three-parameter Weibull distribution fitting. *Granul. Matter* **18**(2), 14 (2016)
- I. Stefanou, J. Sulem, I. Vardoulakis, Three-dimensional Cosserat homogenization of masonry structures: elasticity. *Acta Geotech.* **3**(1), 71–83 (2008)
- I. Stefanou, J. Sulem, I. Vardoulakis, Homogenization of interlocking masonry structures using a generalized differential expansion technique. *Int. J. Solids Struct.* **47**(11–12), 1522–1536 (2010)
- J. Sulem, V. Famin, Thermal decomposition of carbonates in fault zones: slip-weakening and temperature-limiting effects. *J. Geophys. Res.* **114**(B3), B03309 (2009)
- J. Sulem, I. Stefanou, Thermal and chemical effects in shear and compaction bands. *Geomech. Energy Environ.* **6**, 4–21 (2016)
- J. Sulem, I. Vardoulakis, Bifurcation analysis of the triaxial test on rock specimens. A theoretical model for shape and size effect. *Acta Mech.* **83**(3–4), 195–212 (1990)
- J. Sulem, I. Stefanou, E. Veveakis, Stability analysis of undrained adiabatic shearing of a rock layer with Cosserat microstructure. *Granul. Matter* **13**(3), 261–268 (2011)
- P. Trovalusci, M. Ostoja-Starzewski, M.L. De Bellis, A. Murralli, Scale-dependent homogenization of random composites as micropolar continua. *Eur. J. Mech. A/Solids* **49**, 396–407 (2015)
- I. Vardoulakis, *Lecture notes on Cosserat continuum mechanics with application to the mechanics of granular media* (NTU, Athens, 2009)
- I. Vardoulakis, J. Sulem, *Bifurcation Analysis in Geomechanics* (Blackie, Glasgow, 1995)
- E. Veveakis, J. Sulem, I. Stefanou, Modeling of fault gouges with Cosserat continuum mechanics: influence of thermal pressurization and chemical decomposition as coseismic weakening mechanisms. *J. Struct. Geol.* **38**, 254–264 (2012)
- E. Veveakis, I. Stefanou, J. Sulem, Failure in shear bands for granular materials: thermo-hydro-chemo-mechanical effects. *Géotech. Lett.* **3**, 31–36 (2013)

-
- W.M. Wang, L.J. Sluys, R. De Borst, Interaction between material length scale and imperfection size for localisation phenomena in viscoplastic media. *Eur. J. Mech. A/Solids* **15**(3), 447–464 (1996)
- A. Zervos, P. Papanastasiou, I. Vardoulakis, Modelling of localisation and scale effect in thick-walled cylinders with gradient elastoplasticity. *Int. J. Solids Struct.* **38**(30–31), 5081–5095 (2001)
- O.C. Zienkiewicz, R.L. Taylor, J.M. Too, Reduced integration technique in general analysis of plates and shells. *Int. J. Numer. Methods Eng.* **3**(2), 275–290 (1971)
- O.C. Zienkiewicz, R.L. Taylor, J.Z. Zhu, *The Finite Element Method* (Elsevier, Oxford, 2013)



Dispersion of Waves in Micromorphic Media and Metamaterials 20

Angela Madeo and Patrizio Neff

Contents

Introduction	714
Band-Gap Metamaterials	716
Band-Gap Mechanical Metamaterials and the Relaxed Micromorphic Model	718
The Fundamental Role of Micro-inertia in Enriched Continuum Models	719
Notations	722
The Relaxed Micromorphic Model	723
Dispersion Analysis	724
Dispersion Curves for the Relaxed Micromorphic Model	727
The Internal Variable Model	729
The Standard Mindlin-Eringen Model	731
The Micromorphic Model with Curvature $\ \text{Div } P\ ^2 + \ \text{Curl } P\ ^2$	733
The Micromorphic Model with Curvature $\ \text{Div } P\ ^2$	734
Conclusions	736
References	738

Abstract

In this contribution we discuss the interest of using enriched continuum models of the micromorphic type for the description of dispersive phenomena in metamaterials. Dispersion is defined as that phenomenon according to which the speed of propagation of elastic waves is not a constant, but depends on the

A. Madeo (✉)

SMS-ID, INSA-Lyon, Université de Lyon, Villeurbanne cedex, Lyon, France

Institut universitaire de France, Paris Cedex 05, Paris, France

e-mail: angela.madeo@insa-lyon.fr

P. Neff

Fakultät für Mathematik, Universität Duisburg-Essen, Essen, Germany

e-mail: patrizio.neff@uni-due.de

wavelength of the traveling wave. In practice, all materials exhibit dispersion if one considers waves with sufficiently small wavelengths, since all materials have a discrete structure when going down at a suitably small scale. Given the discrete substructure of matter, it is easy to understand that the material properties vary when varying the scale at which the material itself is observed. It is hence not astonishing that the speed of propagation of waves changes as well when considering waves with smaller wavelengths.

In an effort directed toward the modeling of dispersion in materials with architected microstructures (metamaterials), different linear-elastic, isotropic, micromorphic models are introduced, and their peculiar dispersive behaviors are discussed by means of the analysis of the associated dispersion curves. The role of different micro-inertias related to both independent and constrained motions of the microstructure is also analyzed. A special focus is given to those metamaterials which have the unusual characteristic of being able to stop the propagation of mechanical waves and which are usually called band-gap metamaterials. We show that, in the considered linear-elastic, isotropic case, the relaxed micromorphic model, recently introduced by the authors, is the only enriched model simultaneously allowing for the description of non-localities and multiple band-gaps in mechanical metamaterials.

Keywords

Dispersion · Microstructure · Metamaterials · Enriched continuum models · Relaxed micromorphic model · Multi-scale modeling · Gradient micro-inertia · Free micro-inertia · Complete band-gaps · Nonlocal effects

AMS 2010 subject classification: 74A10 (stress), 74A30 (non-simple materials), 74A60 (micro-mechanical theories), 74E15 (crystalline structure), 74M25 (micro-mechanics), 74Q15 (effective constitutive equations)

Introduction

The study of the dispersive behaviors of materials with respect to wave propagation is a central issue in modern mechanics. Dispersion is defined as that phenomenon for which the speed of propagation of waves in a given material changes when changing the wavelength (or, equivalently, the frequency) of the traveling wave (see, e.g., Achenbach 1973). This is a phenomenon which is observed in practically all materials as far as the wavelength of the traveling wave is small enough to interact with the heterogeneities of the material at smaller scales. Indeed, anyone knows that all materials are actually heterogeneous if considering sufficiently small scales: it suffices to go down to the scale of molecules or atoms to be aware of the discrete side of matter. It is hence not astonishing that the mechanical properties of materials are different when considering different scales and that such differences are reflected on the speed of propagation of waves.

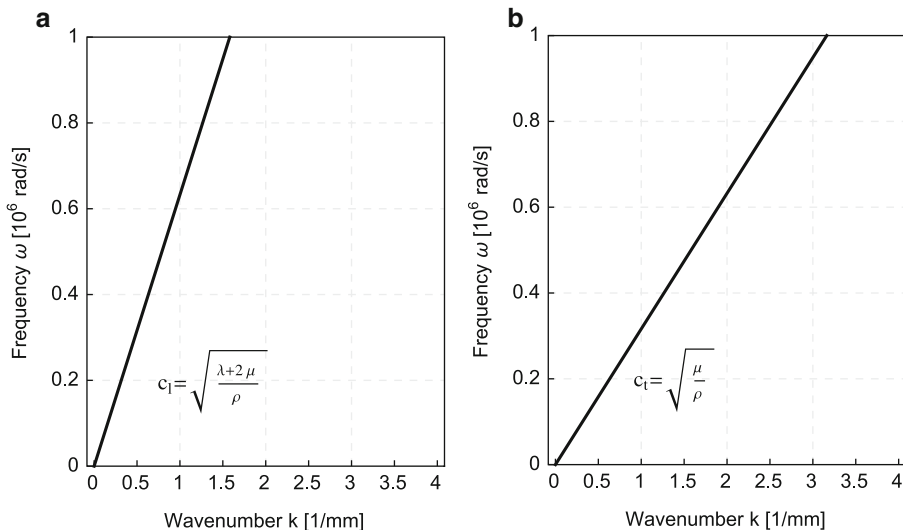


Fig. 1 Typical dispersion curves for longitudinal (a) and transverse (b) waves in (linear-elastic) Cauchy continua (λ and μ are the classical Lamé parameters and ρ is the average mass density). The slopes c_l and c_t are the phase velocities for longitudinal and transverse waves, respectively

The standard approach for the description of the mechanical behavior of materials at large scales is that of using classical Cauchy models which, by their intrinsic nature, are only capable to provide constant speeds of propagation (nondispersive behavior) depending on the elastic moduli of the material and on its mass density. Figure 1 shows the typical behavior of the dispersion curves for Cauchy continua (isotropic linear elasticity). The slope ω/k of the straight lines is called *phase velocity* and is a measure of the speed of propagation of waves in the considered continuum. As a matter of fact, this approach is very effective if one is not interested in the study of wave propagation at high frequencies, since dispersive behaviors are not activated when the wavelength is so large that the traveling wave cannot detect the presence of the underlying microstructure.

On the other hand, Cauchy continuum models are not adapted in all those situations in which dispersive behaviors cannot be neglected. This is the case, for example, when considering so-called *metamaterials*, i.e., materials with architected microstructures which show dispersive behaviors at relatively low frequencies (large wavelengths). Metamaterials are man-made artifacts conceived arranging together small structural elements, usually in periodic or quasiperiodic patterns, in such a way that the resulting material possesses new unorthodox mechanical properties that are not shared by any classical material (see, e.g., Fig. 2). Since the characteristic sizes of microstructures which are usually encountered in mechanical metamaterials typically vary from microns to centimeters, it is not astonishing that such materials exhibit microstructure-related dispersive behaviors for wavelengths which are by far larger than those needed to unveil dispersion in more classical materials. In fact, it is not necessary to have waves with wavelengths that go down

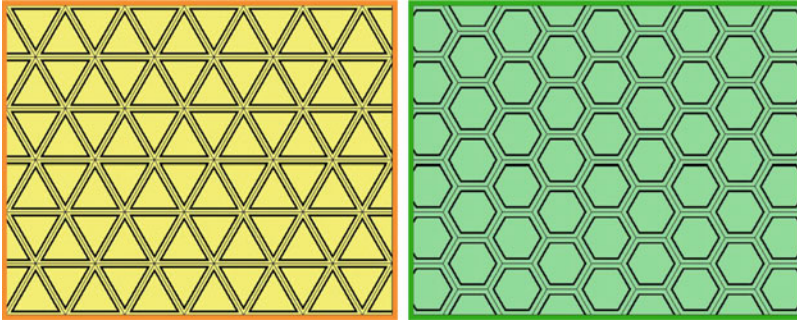


Fig. 2 Examples of periodic microstructures of metamaterials that are susceptible to unveil dispersive behaviors

at the scale of the crystals or of the molecules to start detecting dispersive behaviors. In conclusion, the phenomenon of dispersion in metamaterials cannot be neglected, even at relatively large scales, and models which are more adapted than the classical Cauchy theory need to be introduced.

In this optic, enriched continuum models (micromorphic, second gradient) have been proposed, already in the sixties by Mindlin (1963, 1964) and Eringen (1966, 1999), as suitable generalizations of Cauchy continuum mechanics to study dispersive behaviors. In particular, micromorphic models have shown their efficacy to model rather complex dynamical behaviors by accounting independently for global macroscopic vibrations of the unit cell, as well as for independent vibrations of the microstructures inside the unit cell itself. As we will show in detail, such models present dispersion patterns which are by far richer than the straight lines provided by classical linear elasticity.

As far as second gradient models are concerned, it must be said that they are able to account for some dispersion (as far as a suitable gradient micro-inertia is introduced), but they definitely do not allow for the description of all possible motions inside the unit cell: only those motions of the microstructure which are directly generated by a macro-deformation of the unit cell can be accounted for by such models. More particularly, the dispersion curves for a second gradient continuum are not so different from those of a Cauchy material presented in Fig. 1, except for the fact that they are not straight lines, thus allowing for the description of some dispersion. Nevertheless, due to the fact that the kinematics remains the same of Cauchy media (only the displacement field is introduced in second gradient theories), only two, slightly dispersive, acoustic curves can be found (see also Madeo et al. 2015, 2016d for more details) in such model. This is of course a strong limitation as far as modern metamaterials with extremely complex architectures are concerned.

Band-Gap Metamaterials

In the last years, advanced metamaterials have been conceived and manufactured which exhibit more and more unorthodox properties with respect to the propagation

of mechanical and electromagnetic waves. Actually, materials which are able to “stop” or “bend” the propagation of waves of light or sound with no energetic cost have recently attracted the attention of scientists given the fact that they could suddenly disclose rapid and unimaginable technological advancements. In particular, metamaterials altering electromagnetic wave propagation are currently making the object of intense scientific research (see, e.g., Man et al. 2013; Armenise et al. 2010; Steurer and Sutter-Widmer 2007) for the potential they may have for immediate exotic applications. In fact, they might be used for rendering aircrafts or other vehicles undetectable to radar or for making objects invisible to the human eye. Many other incredible applications have been imagined for such metamaterials such as so-called super-lenses that would allow the human eye to see single viruses or nano-organisms. Notwithstanding the interest raised by such “electromagnetic metamaterials,” they will not make the object of the present work which will be instead centered on the study of “mechanical metamaterials.” As far as metamaterials interacting with mechanical waves are concerned, regrettably, the excitement directed toward their potential applications appears to be less intense than in the previous case, even if their possible fascinating effects on our life are only limited by our imagination. In fact, metamaterials altering elastic wave propagation could be used to build structures absorbing the higher frequencies of earthquake energy, or to conceive naval, automotive, and aeronautical vehicles that are able to absorb external solicitations and shocks thereby drastically improving their internal comfort. What’s more, civil engineering structures which are built in the vicinity of sources of vibrations such as metro lines, tramways, train stations, and so forth would take advantage of the use of these metamaterials to ameliorate the enjoyment of internal and external environments. Based on the same principle, passive engineering devices perfectly able to insulate from noise could be easily conceived and produced at relatively low costs. The conception of waveguides used to optimize energy transfers by collecting waves in slabs or wires, as well as the design of wave screens employed to protect from any sort of mechanical wave, could also experience a new technological revolution. And many other unprecedented applications that, at this juncture, we cannot even envision could be abruptly disclosed once such metamaterials would become easily accessible.

In this contribution, we will focus our attention on those metamaterials which are able to “stop” wave propagation, i.e., metamaterials in which waves within precise frequency ranges cannot propagate. Such frequency intervals at which wave inhibition occurs are known as frequency band-gaps, and their intrinsic characteristics (characteristic values of the gap frequency, extension of the band-gap, etc.) strongly depend on the metamaterial microstructure. Such unorthodox dynamical behavior can be related to two main different phenomena occurring at the microlevel:

- Local resonance phenomena (Mie resonance): The microstructural components, excited at particular frequencies, start oscillating independently of the matrix so capturing the energy of the propagating wave which remains confined at the level of the microstructure. Macroscopic wave propagation thus results to be inhibited.

- Micro-diffusion phenomena (Bragg scattering): When the propagating wave has wavelengths which are small enough to start interacting with the microstructure of the material, reflection and transmission phenomena occur at the micro-level that globally result in an inhibited macroscopic wave propagation.

Such resonance and micro-diffusion mechanisms (usually a mix of the two) are at the basis of both electromagnetic and elastic band-gaps (see, e.g., Steurer and Sutter-Widmer 2007), and they are manifestly related to the particular microstructural topologies of the considered metamaterials. Indeed, it is well known (see, e.g., Steurer and Sutter-Widmer 2007; Armenise et al. 2010; Man et al. 2013) that the characteristics of the microstructures strongly influence the macroscopic band-gap behavior. As it has been said, we will be focused on mechanical waves, even if some of the used theoretical tools can be thought to be suitably generalized for the modeling of electromagnetic waves as well. Such generalizations could open new long-term research directions, for example, in view of the modeling of so-called phoxonic crystals which are simultaneously able to stop both electromagnetic and elastic wave propagation.

Band-Gap Mechanical Metamaterials and the Relaxed Micromorphic Model

As we have previously pointed out, enriched continuum models are well adapted to describe dispersion in metamaterials as well as specific motions of the microstructures inside the unit cells. Nevertheless, not all such enriched models are effective for the description of very complex metamaterials such as those exhibiting band-gap behaviors. As a matter of fact, it has been proved in a series of recent papers Madeo et al. (2014, 2015, 2016a,c,d) that the relaxed micromorphic model is the only enriched continuum model which is simultaneously able to account for:

- The description of the first (and sometimes the second) band-gap in mechanical metamaterials
- The presence of non-localities which are an intrinsic characteristic of microstructured materials, especially when high contrasts in the mechanical properties are present at the level of the microstructure.

Indeed, as we will show in the remainder of this contribution, the relaxed micromorphic model is the most suitable enriched continuum model to be used to simultaneously account for band-gaps and non-localities in mechanical metamaterials. A particular subclass of micromorphic models, the so-called internal variable models, allows for a simplified description of band-gaps, but are not able to account for non-localities related to the microscopic heterogeneities of metamaterials. Such models can be adapted for the description of a particular subclass of metamaterials that are conceived on the basis of the hypothesis of “separation of scales” which basically assumes that all the motions of the microstructure are confined inside the

unit cell and do not have interactions with the motion of the adjacent cells (see, e.g., Pham et al. 2013; Sridhar et al. 2016).

It has to be underlined again that enriched continuum models of the type discussed in this paper are “macroscopic” models in the sense that they allow for the description of the averaged mechanical behavior of metamaterials with complex microstructures while remaining in the simplified framework of continuum mechanics. The main interest of using macroscopic theories for metamaterials can be found in the fact that they feature the introduction of only few parameters which are, in an averaged sense, reminiscent of the presence of the underlying microstructure. If, on the one hand, this fact provides a drastic modeling simplification which is optimal to proceed toward (meta-)structural design, some drawbacks can be reported which are mainly connected to the difficulty of directly relating the introduced macroscopic parameters to the specific characteristics of the microstructure (topology, microstructural mechanical properties, etc.). This difficulty is often seen as a limitation for the effective application of enriched continuum models. As a matter of fact, it is the authors’ belief that such models are a necessary step if one wants to proceed toward the engineering design of *metastructures*, i.e., structures which are made up of metamaterials as building blocks. Of course, the proposed model will introduce a certain degree of simplification, but it is exactly this simplicity that makes possible to envision the following step which is that of proceeding toward the design of complex structures made of metamaterials.

To be more precise, as we will see, the relaxed micromorphic model is able to describe the onset of the first (and sometimes the second) band-gap in mechanical metamaterials. In order to catch more complex behaviors, the kinematics and the constitutive relations of the proposed model should be further enriched in a way that is not yet completely clear. Nevertheless, we do not see this fact as a true limitation since we intend to use the unorthodox dynamical behavior of some metamaterials exhibiting band-gaps to fit, by inverse approach, the parameters of the relaxed micromorphic model following what has been done, e.g., in Madeo et al. (2016a). This fitting, when successfully concluded for a certain number of specific metamaterials, will allow the design of metastructures by means of tools which are familiar to engineers, such as finite element codes.

Of course, as classical Cauchy models show their limits for the description of the dynamical behavior of metamaterials, even at low frequencies, the relaxed micromorphic model will show its limit for higher frequencies, yet remaining accurate enough to account for some macroscopic manifestations of microstructure.

The Fundamental Role of Micro-inertia in Enriched Continuum Models

As far as enriched continuum models are concerned, a central issue which is also an open scientific question is that of identifying the role of so-called micro-inertia terms on the dispersive behavior of such media. As a matter of fact, enriched continuum models usually provide a richer kinematics, with respect to the

classical macroscopic displacement field alone, which is related to the possibility of describing the motions of the microstructure inside the unit cell. The adoption of such enriched kinematics (given by the displacement field u and the micro-distortion tensor P ; see, e.g., Mindlin 1964; Eringen 1999; Ghiba et al. 2014; Madeo et al. 2014, 2015, 2016a,c,d; Neff et al. 2014b), as we will see, allows for the introduction of constitutive laws for the strain energy density that are able to describe the mechanical behavior of some metamaterials in the static regime. When the dynamical regime is considered, things become even more delicate since the choice of micro-inertia terms to be introduced in the kinetic energy density must be carefully based on:

- A compatibility with the chosen kinematics and constitutive laws used for the description of the static regime,
- The specific inertial characteristics of the metamaterial that one wants to describe (e.g., eventual coupling of the motion of the microstructure with the macro motions of the unit cell, specific resistance of the microstructure to independent motion, etc.).

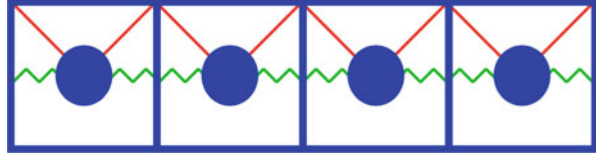
It has to be explicitly mentioned that, in this contribution, we voluntarily limit ourselves to consider the linear-elastic, isotropic case with the aim of unveiling the most fundamental properties of enriched continuum models for the description of metamaterials in the dynamic regime. As a matter of fact, this simplification will allow us to understand the essence of each introduced term and its effect on the description of the mechanical behavior of particular metamaterials. Indeed, enriched continuum models of the micromorphic type featuring the description of band-gaps when introducing nonlinearities in the micro-inertia terms can also be found in the literature Stefano et al. (2011), but the interpretation of the introduced nonlinearities would be more complex to be undertaken on a phenomenological ground.

In the present contribution, for all the models that will be presented (relaxed micromorphic, Mindlin-Eringen micromorphic, internal variable, etc.), we will consider a kinetic energy of the type

$$J = \underbrace{\frac{1}{2}\rho \|u_{,i}\|^2}_{\text{Cauchy inertia free micro-inertia}} + \underbrace{\frac{1}{2}\eta \|P_{,i}\|^2 + \frac{1}{2}\bar{\eta}_1 \|\text{dev sym } \nabla u_{,i}\|^2 + \frac{1}{2}\bar{\eta}_2 \|\text{skew } \nabla u_{,i}\|^2 + \frac{1}{6}\bar{\eta}_3 \text{tr}(\nabla u_{,i})^2}_{\text{gradient micro-inertia}}, \quad (1)$$

where ρ is the value of the average macroscopic mass density of the considered metamaterial, η is the free micro-inertia density, and the $\bar{\eta}_i, i = \{1, 2, 3\}$ are the gradient micro-inertia densities associated to the different terms of the Cartan-Lie decomposition of ∇u . The comprehension of the terms appearing in such expression of the kinetic energy density is fundamental for the correct understanding and exploitation of enriched continuum models for the description of the mechanical behavior of metamaterials in the dynamic regime.

Fig. 3 Schematic representation of a system showing free and gradient micro-inertia



For this reason we discuss separately each term appearing in Eq. (1) by giving a general interpretation of its effect on the dynamics of metamaterials:

- The Cauchy inertia term $\frac{1}{2}\rho \|u_{,t}\|^2$ is the macroscopic inertia introduced in classical linear elasticity. It allows to describe the vibrations associated to the macroscopic displacement field. In an enriched continuum mechanical modeling framework, this means that such terms account for the inertia to vibration of the unit cells considered as material points (or representative volume elements) with apparent mass density ρ .
- The term $\frac{1}{2}\eta \|P_{,t}\|^2$ accounts for the inertia of the microstructure alone: with reference to Fig. 3, the free micro-inertia η can be thought to be due to the green springs connecting the blue masses inside the unit cells. The higher is the value of η , the stiffer we can imagine the green springs. We called η *free* micro-inertia (Madeo et al. 2016b) since it represents the inertia of the microstructure seen as a micro-system whose vibration can be independent of the vibration of the unit cells. An inertia term of this type is mandatory whenever one considers an enriched model of the micromorphic type, i.e., a model that features an enriched kinematics (u, P) . Indeed, it would be senseless to introduce an enriched kinematics and an enriched constitutive form for the strain energy density and then avoid to introduce this free micro-inertia in the model. It would be like introducing a complex constitutive structure to describe in detail the mechanical behavior of microstructured materials while not giving to the model the possibility of activating the vibrations of such microstructures. The free micro-inertia allows us to account for the vibrations of the microstructures that typically appear for high frequencies (i.e., small wavelengths comparable with the characteristic size of the microstructure) in a huge variety of mechanical metamaterials. We explicitly mention the fact that the free micro-inertia term could take advantage of the Cartan-Lie decomposition of the micro-distortion tensor P in its dev sym (trace-free symmetric-), skew (skew-symmetric-) and tr (trace-) part as it is done for the following term. In this work we do not explicitly consider this possibility, even if we will evoke some of the effects of such a choice on the dispersion curves later on.
- The gradient micro-inertia term is of the type $\bar{\eta} \|\nabla u_{,t}\|^2$, and, when split using a Cartan-Lie decomposition, it takes the form shown in Eq. (1) (Madeo et al. 2016b). Such term allows to account for some specific vibrations of the microstructure which are directly related to the deformation of the unit cell at the higher scale. With reference to Fig. 3, the term $\bar{\eta}$ can be related to the effect that the red bars have on the motion of the blue masses. Increasing the value

of $\bar{\eta}$ is tantamount to make the red bars stiffer, so that we can finally say that a nonvanishing $\bar{\eta}$ constrains the motions of the microstructure to be somehow related to the macroscopic deformation of the unit cell. In fact, we can easily imagine that a deformation of the unit cell can induce a motion of the mass through the motion of the red bars. With respect to a similar metamaterial in which only the free micro-inertia η is present, we can imagine that a metamaterial with a nonvanishing gradient micro-inertia $\bar{\eta}$ is more “local” in the sense that the vibration of the microstructure is more confined to the unit cell than if only the green springs were present. This gradient micro-inertia term brings additional information with respect to the free micro-inertia term previously described, and this is translated in the behavior of some dispersion curves that, as we will see, can be flattened when increasing the value of $\bar{\eta}_1, \bar{\eta}_2$ or $\bar{\eta}_3$.

Notations

In this contribution, we denote by $\mathbb{R}^{3 \times 3}$ the set of real 3×3 second-order tensors, written with capital letters. We denote respectively by $\cdot, :$ and $\langle \cdot, \cdot \rangle$ a simple and double contraction and the scalar product between two tensors of any suitable order. (For example, $(A \cdot v)_i = A_{ij} v_j, (A \cdot B)_{ik} = A_{ij} B_{jk}, A : B = A_{ij} B_{ji}, (C \cdot B)_{ijk} = C_{ijp} B_{pk}, (C : B)_i = C_{ijp} B_{pj}, \langle v, w \rangle = v \cdot w = v_i w_i, \langle A, B \rangle = A_{ij} B_{ij}$ etc.) Everywhere we adopt the Einstein convention of sum over repeated indices if not differently specified. The standard Euclidean scalar product on $\mathbb{R}^{3 \times 3}$ is given by $\langle X, Y \rangle_{\mathbb{R}^{3 \times 3}} = \text{tr}(X \cdot Y^T)$, and thus the Frobenius tensor norm is $\|X\|^2 = \langle X, X \rangle_{\mathbb{R}^{3 \times 3}}$. In the following we omit the index $\mathbb{R}^3, \mathbb{R}^{3 \times 3}$. The identity tensor on $\mathbb{R}^{3 \times 3}$ will be denoted by $\mathbb{1}$, so that $\text{tr}(X) = \langle X, \mathbb{1} \rangle$.

We consider a body which occupies a bounded open set B_L of the three-dimensional Euclidian space \mathbb{R}^3 and assume that its boundary ∂B_L is a smooth surface of class C^2 . An elastic material fills the domain $B_L \subset \mathbb{R}^3$, and we refer the motion of the body to rectangular axes Ox_i .

For vector fields v with components in $H^1(B_L)$, i.e., $v = (v_1, v_2, v_3)^T, v_i \in H^1(B_L)$, we define $\nabla v = ((\nabla v_1)^T, (\nabla v_2)^T, (\nabla v_3)^T)^T$, while for tensor fields P with rows in $H(\text{curl}; B_L)$, resp. $H(\text{div}; B_L)$, i.e., $P = (P_1^T, P_2^T, P_3^T)^T, P_i \in H(\text{curl}; B_L)$ resp. $P_i \in H(\text{div}; B_L)$ we define $\text{Curl } P = ((\text{curl } P_1)^T, (\text{curl } P_2)^T, (\text{curl } P_3)^T)^T$, $\text{Div } P = (\text{div } P_1, \text{div } P_2, \text{div } P_3)^T$.

As for the kinematics of the considered micromorphic continua, we introduce the functions

$$\chi(X, t) : B_L \rightarrow \mathbb{R}^3, \quad P(X, t) : B_L \rightarrow \mathbb{R}^{3 \times 3},$$

which are known as *placement* vector field and *micro-distortion* tensor, respectively. The physical meaning of the placement field is that of locating, at any instant t , the current position of the material particle $X \in B_L$, while the micro-distortion field describes deformations of the microstructure embedded in the material particle X

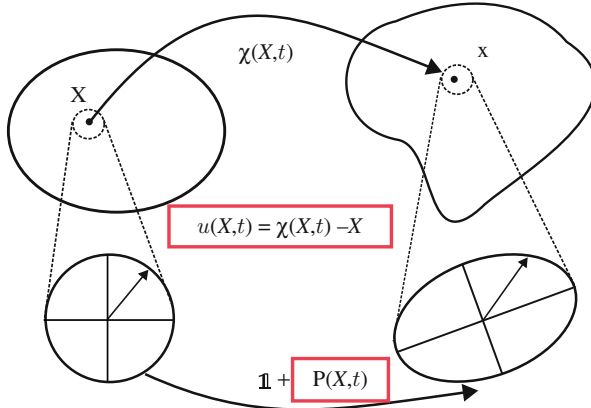


Fig. 4 Schematic representation of the enriched kinematics describing the motion of enriched models of the micromorphic type. A model of this type has 3 (components of u) + 9 (components of P) = 12 degrees of freedom (DOF)

(see Fig. 4). As it is usual in continuum mechanics, the displacement field can also be introduced as the function $u(X, t) : B_L \rightarrow \mathbb{R}^3$ defined as

$$u(X, t) = \chi(X, t) - X.$$

The Relaxed Micromorphic Model

The relaxed micromorphic model endows Mindlin-Eringen’s representation with the second-order **dislocation density tensor** $\alpha = -\text{Curl}P$ instead of the full gradient ∇P . (The dislocation tensor is defined as $\alpha_{ij} = -(\text{Curl}P)_{ij} = -P_{ih,k}\epsilon_{jkh}$, where ϵ is the Levi-Civita tensor and Einstein notation of sum over repeated indices is used.) In the isotropic case, the energy of the relaxed micromorphic model reads (Ghiba et al. 2014; Madeo et al. 2014, 2015, 2016c; Neff et al. 2014a,b):

$$\begin{aligned}
 W = & \underbrace{\mu_e \|\text{sym}(\nabla u - P)\|^2 + \frac{\lambda_e}{2} (\text{tr}(\nabla u - P))^2}_{\text{isotropic elastic-energy}} + \underbrace{\mu_c \|\text{skew}(\nabla u - P)\|^2}_{\text{rotational elastic coupling}} \\
 & + \underbrace{\mu_{\text{micro}} \|\text{sym} P\|^2 + \frac{\lambda_{\text{micro}}}{2} (\text{tr}P)^2}_{\text{micro-self-energy}} + \underbrace{\frac{\mu_e L_c^2}{2} \|\text{Curl}P\|^2}_{\text{isotropic curvature}}, \tag{2}
 \end{aligned}$$

where all the introduced constitutive parameters are assumed to be constant. The model is well-posed in the static and dynamical case including when $\mu_c = 0$; see Neff et al. (2014a) and Ghiba et al. (2014).

The relaxed micromorphic model counts 6 constitutive parameters in the isotropic case ($\mu_e, \lambda_e, \mu_{\text{micro}}, \lambda_{\text{micro}}, \mu_c, L_c$). The characteristic length L_c is

intrinsically related to nonlocal effects due to the fact that it weights a suitable combination of first-order space derivatives in the strain energy density (2). For a general presentation of the features of the relaxed micromorphic model in the anisotropic setting, we refer to Barbagallo et al. (2016).

The associated equations of motion in strong form, obtained by a classical least action principle, take the form (see Madeo et al. (2014, 2015, 2016d) and Neff et al. 2014a)

$$\rho u_{,tt} - \text{Div}[\mathcal{I}] = \text{Div}[\tilde{\sigma}], \quad \eta P_{,tt} = \tilde{\sigma} - s - \text{Curl} m, \quad (3)$$

where

$$\begin{aligned} \mathcal{I} &= \bar{\eta}_1 \text{dev sym } \nabla u_{,tt} + \bar{\eta}_2 \text{skew } \nabla u_{,tt} + \frac{1}{3} \bar{\eta}_3 \text{tr}(\nabla u_{,tt}), \\ \tilde{\sigma} &= 2\mu_e \text{sym}(\nabla u - P) + \lambda_e \text{tr}(\nabla u - P) \mathbb{1} + 2\mu_c \text{skew}(\nabla u - P), \quad (4) \\ s &= 2\mu_{\text{micro}} \text{sym} P + \lambda_{\text{micro}} \text{tr}(P) \mathbb{1}, \\ m &= \mu_e L_c^2 \text{Curl} P. \end{aligned}$$

The fact of adding a gradient micro-inertia in the kinetic energy (1) modifies the strong-form PDEs of the relaxed micromorphic model with the addition of the new term \mathcal{I} (Madeo et al. 2016b). Of course, boundary conditions would also be modified with respect to the ones presented in Madeo et al. (2016a,d), but they will not be the object of study of the present contribution.

Dispersion Analysis

We rapidly recall in this subsection how, starting from the equations of motion in strong form for the relaxed micromorphic medium, it is possible to obtain the associated dispersion curves by following standard techniques. We start by making a plane-wave ansatz which means that we assume that all the 12 scalar components of the unknown fields (In what follows, we will not differentiate anymore the Lagrangian space variable X and the Eulerian one x . In general, such undifferentiated space variable will be denoted as $x = (x_1, x_2, x_3)^T$) $u(x, t)$ and $P(x, t)$ only depend on the component x_1 of the space variable x which is also assumed to be the direction of the traveling wave. With this unique assumption, together with the introduction of the new variables

$$\begin{aligned} P^S &:= \frac{1}{3} \text{tr}(P), & P_{[ij]} &:= (\text{skew} P)_{ij} = \frac{1}{2} (P_{ij} - P_{ji}), \quad (5) \\ P^D &:= P_{11} - P^S, & P_{(ij)} &:= (\text{sym} P)_{ij} = \frac{1}{2} (P_{ij} + P_{ji}), \\ P^V &:= P_{22} - P_{33}, & i, j &= \{1, 2, 3\}, \end{aligned}$$

the equations of motions (3) can be simplified and rewritten, after suitable manipulation, as (see d'Agostino et al. (2016) and Madeo et al. (2014, 2015, 2016c) for additional details):

- A set of three equations only involving longitudinal quantities:

$$\begin{aligned} \rho \ddot{u}_1 \frac{2\bar{\eta}_1 + \bar{\eta}_3}{3} \ddot{u}_{1,11} &= (2\mu_e + \lambda_e) u_{1,11} - 2\mu_e P_{,1}^D - (2\mu_e + 3\lambda_e) P_{,1}^S, \\ \eta \ddot{P}^D &= \frac{4}{3} \mu_e u_{1,1} + \frac{1}{3} \mu_e L_c^2 P_{,11}^D - \frac{2}{3} \mu_e L_c^2 P_{,11}^S - 2(\mu_e + \mu_{\text{micro}}) P^D, \\ \eta \ddot{P}^S &= \frac{2\mu_e + 3\lambda_e}{3} u_{1,1} - \frac{1}{3} \mu_e L_c^2 P_{,11}^D + \frac{2}{3} \mu_e L_c^2 P_{,11}^S \\ &\quad - (2\mu_e + 3\lambda_e + 2\mu_{\text{micro}} + 3\lambda_{\text{micro}}) P^S, \end{aligned} \quad (6)$$

- Two sets of three equations only involving transverse quantities in the ξ -th direction, with $\xi = 2, 3$:

$$\begin{aligned} \rho \ddot{u}_\xi \frac{\bar{\eta}_1 + \bar{\eta}_2}{2} \ddot{u}_{\xi,11} &= (\mu_e + \mu_c) u_{\xi,11} - 2\mu_e P_{(1\xi),1} + 2\mu_c P_{[1\xi],1}, \\ \eta \ddot{P}_{(1\xi)} &= \mu_e u_{\xi,1} + \frac{1}{2} \mu_e L_c^2 P_{(1\xi),11} + \frac{1}{2} \mu_e L_c^2 P_{[1\xi],11} \\ &\quad - 2(\mu_e + \mu_{\text{micro}}) P_{(1\xi)}, \\ \eta \ddot{P}_{[1\xi]} &= -\mu_c u_{\xi,1} + \frac{1}{2} \mu_e L_c^2 P_{(1\xi),11} + \frac{1}{2} \mu_e L_c^2 P_{[1\xi],11} - 2\mu_c P_{[1\xi]}, \end{aligned} \quad (7)$$

- One equation only involving the variable $P_{(23)}$:

$$\eta \ddot{P}_{(23)} = -2(\mu_e + \mu_{\text{micro}}) P_{(23)} + \mu_e L_c^2 P_{(23),11}, \quad (8)$$

- One equation only involving the variable $P_{[23]}$:

$$\eta \ddot{P}_{[23]} = -2\mu_c P_{[23]} + \mu_e L_c^2 P_{[23],11},$$

- One equation only involving the variable P^V :

$$\eta \ddot{P}^V = -2(\mu_e + \mu_{\text{micro}}) P^V + \mu_e L_c^2 P_{,11}^V. \quad (9)$$

Once that this simplified form of the equations of motion is obtained, we look for a wave-form solution of the type

$$\begin{aligned}
 v_1 &= \beta_1 e^{i(kx_1 - \omega t)}, & v_2 &= \beta_2 e^{i(kx_1 - \omega t)}, & v_3 &= \beta_3 e^{i(kx_1 - \omega t)}, \\
 v_4 &= \beta_4 e^{i(kx_1 - \omega t)}, & v_5 &= \beta_5 e^{i(kx_1 - \omega t)}, & v_6 &= \beta_6 e^{i(kx_1 - \omega t)},
 \end{aligned}
 \tag{10}$$

where $\beta_1, \beta_2, \beta_3 \in \mathbb{C}^3$ and $\beta_4, \beta_5, \beta_6 \in \mathbb{C}$ are the unknown amplitudes of the considered waves, ω is the frequency, and k is the wavenumber and where, for compactness, we set

$$\begin{aligned}
 v_1 &:= (u_1, P_{11}^D, P^S), & v_2 &:= (u_2, P_{(12)}, P_{[12]}), & v_3 &:= (u_3, P_{(13)}, P_{[13]}), \\
 v_4 &:= P_{(23)}, & v_5 &:= P_{[23]}, & v_6 &:= P^V.
 \end{aligned}
 \tag{11}$$

Replacing the wave-form (10) and (11) in the equations of motion (6), (7), (8) and (9) and simplifying, we end up with the following systems of algebraic equations:

$$A_1(\omega, k) \cdot \beta_1 = 0, \quad A_\tau(\omega, k) \cdot \beta_\tau = 0, \quad \tau = 2, 3, \quad A_4(\omega, k) \cdot \alpha = 0,
 \tag{12}$$

where we set $\alpha = (\beta_4, \beta_5, \beta_6)$ and

$$\begin{aligned}
 A_1(\omega, k) &= \begin{pmatrix} -\omega^2 \left(1 + k^2 \frac{2\bar{\eta}_1 + \bar{\eta}_3}{3\rho}\right) + c_p^2 k^2 & i k 2 \mu_e / \rho & i k (2 \mu_e + 3 \lambda_e) / \rho \\ -i k \frac{4}{3} \mu_e / \eta & -\omega^2 + \frac{1}{3} k^2 c_m^2 + \omega_s^2 & -\frac{2}{3} k^2 c_m^2 \\ -\frac{1}{3} i k (2 \mu_e + 3 \lambda_e) / \eta & -\frac{1}{3} k^2 c_m^2 & -\omega^2 + \frac{2}{3} k^2 c_m^2 + \omega_p^2 \end{pmatrix}, \\
 A_2(\omega, k) &= A_3(\omega, k) \\
 &= \begin{pmatrix} -\omega^2 \left(1 + k^2 \frac{\bar{\eta}_1 + \bar{\eta}_2}{2\rho}\right) + k^2 c_s^2 & i k 2 \mu_e / \rho & -i \frac{k}{\rho} \eta \omega_r^2, \\ -i k \mu_e / \eta, & -\omega^2 + \frac{c_m^2}{2} k^2 + \omega_s^2 & \frac{c_m^2}{2} k^2 \\ \frac{i}{2} \omega_r^2 k & \frac{c_m^2}{2} k^2 & -\omega^2 + \frac{c_m^2}{2} k^2 + \omega_r^2 \end{pmatrix}, \\
 A_4(\omega, k) &= \begin{pmatrix} -\omega^2 + c_m^2 k^2 + \omega_s^2 & 0 & 0 \\ 0 & -\omega^2 + c_m^2 k^2 + \omega_r^2 & 0 \\ 0 & 0 & -\omega^2 + c_m^2 k^2 + \omega_s^2 \end{pmatrix}.
 \end{aligned}
 \tag{13}$$

In the definition of the matrices $A_i, i = \{1, 2, 3, 4\}$, the following characteristic quantities have also been introduced:

$$\begin{aligned} \omega_s &= \sqrt{\frac{2(\mu_e + \mu_{\text{micro}})}{\eta}}, & \omega_r &= \sqrt{\frac{2\mu_c}{\eta}}, & \omega_p &= \sqrt{\frac{(3\lambda_e + 2\mu_e) + (3\lambda_{\text{micro}} + 2\mu_{\text{micro}})}{\eta}}, \\ c_m &= \sqrt{\frac{\mu_e L_c^2}{\eta}}, & c_p &= \sqrt{\frac{\lambda_e + 2\mu_e}{\rho}}, & c_s &= \sqrt{\frac{\mu_e + \mu_c}{\rho}}. \end{aligned} \quad (14)$$

Dispersion Curves for the Relaxed Micromorphic Model

The dispersion curves for the relaxed micromorphic model are the solutions $\omega = \omega(k)$ of the algebraic equations

$$\underbrace{\det A_1(\omega, k) = 0}_{\text{longitudinal}}, \quad \underbrace{\det A_2(\omega, k) = \det A_3(\omega, k) = 0}_{\text{transverse}}, \quad \underbrace{\det A_4(\omega, k) = 0}_{\text{uncoupled}}. \quad (15)$$

In order to plot such curves, we chose the values of the constitutive parameters to be those in Table 1.

These numerical values will be also kept for the simulations relative to the other models in the subsequent sections if not differently specified. (We explicitly remark that the values of the elastic coefficients must be chosen in such a way to respect the conditions of positive definiteness of the strain energy density (Madeo et al. 2014, 2015), i.e., $\mu_e > 0$, $\mu_{\text{micro}} > 0$, $3\lambda_e + 2\mu_e > 0$, $3\lambda_{\text{micro}} + 2\mu_{\text{micro}} > 0$, $\mu_c \geq 0$). We explicitly remark that the chosen values correspond to a relatively soft material, but completely analogous results can be obtained with any other combination of the parameters with the only constraint of respecting positive definiteness of the strain energy density.

We start by presenting in Fig. 5 the dispersion curves for the relaxed micromorphic model obtained with vanishing gradient micro-inertia $\bar{\eta}$. In the figures presented thereafter, we adopt the following acronyms:

Table 1 Values of the parameters used in the numerical simulations (left) and corresponding values of the Lamé parameters and of the Young modulus and Poisson ratio (right). For the formulas needed to calculate the homogenized macroscopic parameters starting from the microscopic ones see Barbagallo et al. (2016).

Parameter	Value	Unit	Parameter	Value	Unit
μ_e	200	MPa	λ_{macro}	82.5	MPa
$\lambda_e = 2\mu_e$	400	MPa	μ_{macro}	66.7	MPa
$\mu_c = 5\mu_e$	1000	MPa	E_{macro}	170	MPa
μ_{micro}	100	MPa			
λ_{micro}	100	MPa			
L_c	1	mm			
ρ	2000	kg/m ³			
η	10 ⁻²	kg/m ³			
$\bar{\eta}$	10 ⁻¹	kg/m ³			

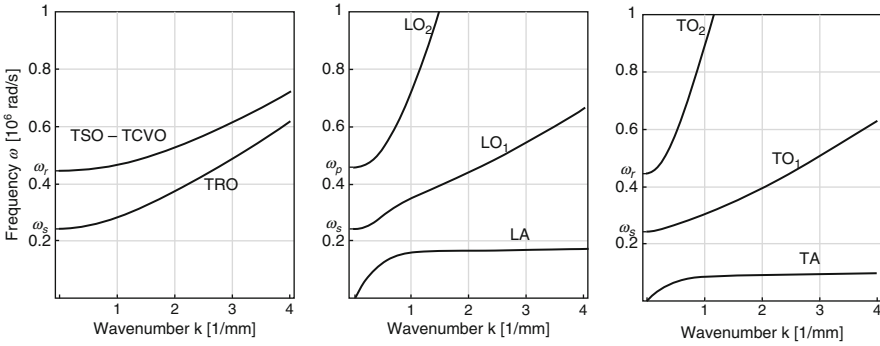


Fig. 5 Dispersion relations $\omega = \omega(k)$ of the **relaxed micromorphic model** for the uncoupled (*left*), longitudinal (*center*), and transverse (*right*) waves with nonvanishing free micro-inertia $\eta \neq 0$ and vanishing gradient micro-inertia $\bar{\eta} = 0$

- TRO: transverse rotational optic,
- TSO: transverse shear optic,
- TCVO: transverse constant-volume optic,
- LA: longitudinal acoustic,
- LO₁-LO₂: 1st and 2nd longitudinal optic,
- TA: transverse acoustic,
- TO₁-TO₂: 1st and 2nd transverse optic.

It can be seen from Fig.5 that the main characteristic feature of the relaxed micromorphic model is that of having a horizontal asymptote for the acoustic waves. It is exactly this feature that allows the relaxed model to account for the description of band-gaps in nonlocal band-gap metamaterials. Indeed, a complete band-gap can be identified in the interval of frequencies individuated by the horizontal asymptote of the longitudinal acoustic wave *LA* and the characteristic frequency ω_s . In such interval, the wavenumber becomes purely imaginary for all the considered waves (longitudinal, transverse, and uncoupled) which means that the wave can actually not propagate.

Now, we show in Fig.6 the results obtained for nonvanishing gradient micro-inertia $\bar{\eta} \neq 0$ (here and in the sequel, we suppose in the expression (1) to have $\bar{\eta}_1 = 0, \bar{\eta}_2 = \bar{\eta}_3 = \bar{\eta}$, if not differently specified). The effect of the addition of such gradient micro-inertia term is that of flattening the optic curves so allowing for the possibility of the onset of a second longitudinal and transverse band-gap if the gradient inertia parameters are sufficiently high. The fact of having split the gradient micro-inertia in three parts allows to flatten the longitudinal and transverse waves separately. Moreover, we remark that the addition of gradient micro-inertiae $\bar{\eta}_1, \bar{\eta}_2$, and $\bar{\eta}_3$ has no effect on the cutoff frequencies ω_s, ω_p , and ω_r which only depend on the free micro-inertia η (and of course on the constitutive parameters). Indeed, if an equivalent Cartan-Lie decomposition of the free micro-inertia term would have been performed, the result would have been that of gaining the independence of the

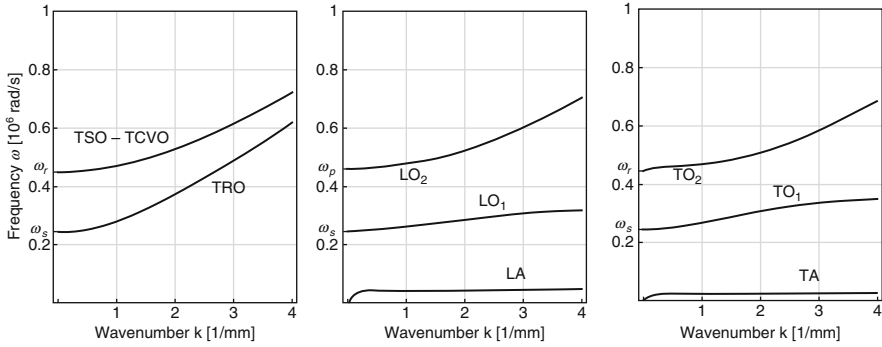


Fig. 6 Dispersion relations $\omega = \omega(k)$ of the **relaxed micromorphic model** for the uncoupled (*left*), longitudinal (*center*), and transverse (*right*) waves with nonvanishing free micro-inertia $\eta \neq 0$ and nonvanishing gradient micro-inertia $\bar{\eta} \neq 0$

three cutoff frequencies which could be then translated independently one from the other along the y -axis.

It is worth to explicitly remark that the observed flattening effect of the micro-inertia $\bar{\eta}$ can be interpreted as the fact that some modes propagate more slowly when $\bar{\eta}$ is nonvanishing. Indeed, it is well known (Achenbach 1973) that the slope of the dispersion curves in a point is a measure of the speed of propagation of the considered waves (phase velocity). With this definition in mind, it is easy to infer by direct comparison of Figs. 5 and 6 that the longitudinal and transverse waves are much slower in the second case, when $\bar{\eta} \neq 0$. On the other hand, we can also remark that the modes corresponding to the uncoupled waves are not affected by the introduction of the gradient micro-inertia $\bar{\eta}$.

With reference to previous works Madeo et al. (2014, 2015, 2016c,d), it is possible to observe that the oblique asymptotes of the waves LO_1 , TO_1 as well as the asymptotes of the uncoupled waves are directly proportional to the value of the characteristic length L_c . This is equivalent to say that such waves propagate, thanks to the non-localities that are intrinsic of the proposed relaxed micromorphic model. When the gradient micro-inertia is introduced, the waves LO_1 , TO_1 are flattened, so that the effect of non-localities results to be reduced. The non-local behavior on the uncoupled waves, instead, is not affected by the introduction of the gradient micro-inertia.

The Internal Variable Model

We recall (see Neff et al. 2014b) that the energy for the internal variable model does not include higher space derivatives of the micro-distortion tensor P and, in the isotropic case, takes the form

$$\begin{aligned}
 W = & \underbrace{\mu_e \|\operatorname{sym}(\nabla u - P)\|^2 + \frac{\lambda_e}{2} (\operatorname{tr}(\nabla u - P))^2}_{\text{isotropic elastic-energy}} + \underbrace{\mu_c \|\operatorname{skew}(\nabla u - P)\|^2}_{\text{rotational elastic coupling}} \\
 & + \underbrace{\mu_{\text{micro}} \|\operatorname{sym} P\|^2 + \frac{\lambda_{\text{micro}}}{2} (\operatorname{tr} P)^2}_{\text{micro-self-energy}},
 \end{aligned} \tag{16}$$

Due to the absence of space derivatives of the micro-distortion tensor P in the strain energy density, such model intrinsically excludes the possibility of modeling non-localities. If, in a first instance, this can be a reasonable simplification for a certain class of metamaterials for which the “separation of scale” hypothesis can be verified (Pham et al. 2013; Sridhar et al. 2016), the fact of neglecting the effect of the deformation of the unit cell on the adjacent ones can be too restrictive, especially when metamaterials with high contrasts of the mechanical properties at the micro-level are considered.

The equations of motion for the internal variable model, obtained as the result of a least action principle based on the strain energy density (16) and the kinetic energy (1), are

$$\rho u_{,tt} - \operatorname{Div}[\mathcal{I}] = \operatorname{Div}[\tilde{\sigma}], \quad \eta P_{,tt} = \tilde{\sigma} - s, \tag{17}$$

where

$$\begin{aligned}
 \mathcal{I} &= \bar{\eta}_1 \operatorname{dev} \operatorname{sym} \nabla u_{,tt} + \bar{\eta}_2 \operatorname{skew} \nabla u_{,tt} + \frac{1}{3} \bar{\eta}_3 \operatorname{tr}(\nabla u_{,tt}), \\
 \tilde{\sigma} &= 2\mu_e \operatorname{sym}(\nabla u - P) + \lambda_e \operatorname{tr}(\nabla u - P) \mathbb{1} + 2\mu_c \operatorname{skew}(\nabla u - P), \tag{18} \\
 s &= 2\mu_{\text{micro}} \operatorname{sym} P + \lambda_{\text{micro}} \operatorname{tr}(P) \mathbb{1}.
 \end{aligned}$$

We present the dispersion relations obtained for the internal variable model with a vanishing gradient inertia (Fig. 7) and for a nonvanishing gradient inertia (Fig. 8). We start by noticing in Fig. 7 that, suitably choosing the value of the cut-off frequencies, the internal variable model allows for the description of up to 2 band-gaps. The pattern of longitudinal and transverse waves, even if not including non-localities, still remains realistic for at least some particular metamaterials. On the other hand, this “absence of non-locality” drastically affects the uncoupled waves which become perfectly horizontal lines (local resonances of some modes inside the unit cell), which can be a too drastic simplification for three-dimensional metamaterials.

Figure 8 shows the effect of the addition of the gradient micro-inertia $\bar{\eta} \|\nabla u_{,t}\|^2$ on the internal variable model. By direct observation of Fig. 8, we can notice that suitably choosing the relative position of ω_r and ω_p , the internal variable model allows to account for 3 band-gaps. We thus have an extra band-gap with respect to

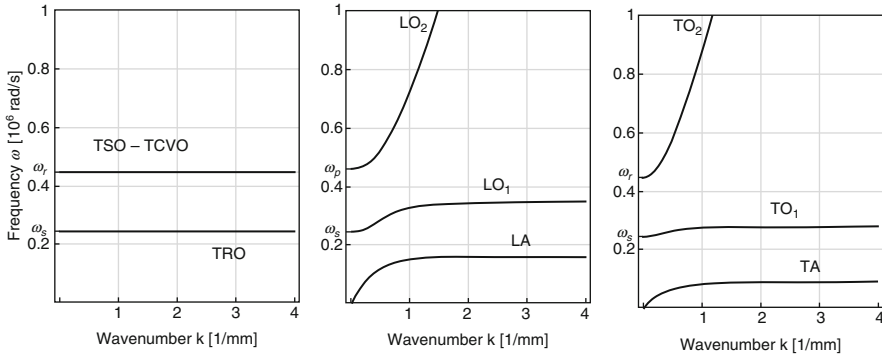


Fig. 7 Dispersion relations $\omega = \omega(k)$ of the **internal variable model** for the uncoupled (*left*), longitudinal (*center*) and transverse (*right*) waves with nonvanishing free micro-inertia $\eta \neq 0$ and vanishing gradient micro-inertia $\bar{\eta} = 0$

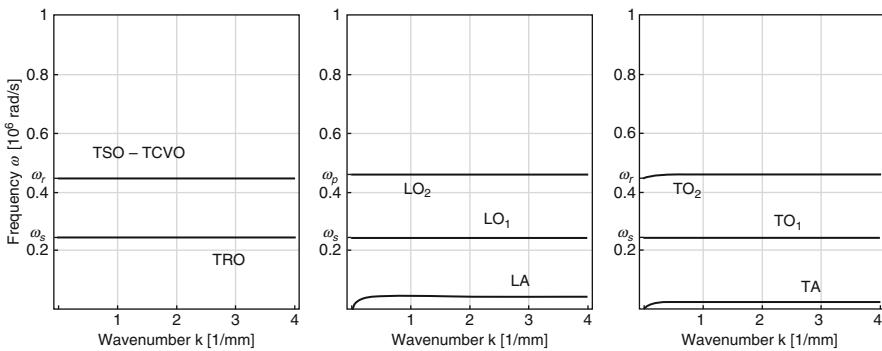


Fig. 8 Dispersion relations $\omega = \omega(k)$ of the **internal variable model** for the uncoupled (*left*), longitudinal (*center*), and transverse (*right*) waves with nonvanishing free micro-inertia $\eta \neq 0$ and nonvanishing gradient micro-inertia $\bar{\eta} \neq 0$

the case with vanishing gradient inertia (Fig. 7) and to the analogous case for the relaxed micromorphic model (see Fig. 6), but we are not able to consider non-local effects. The characteristic behavior of the dispersion curves becomes unrealistic in this case, given that almost only local resonances (dispersion curves which become horizontal straight lines) are provided. It turns out that the fact of excluding the possibility of describing non-local effects in metamaterials becomes too restrictive and unphysical for the description of the great majority of metamaterials.

The Standard Mindlin-Eringen Model

In this section we discuss the effect on the Mindlin-Eringen model of the addition of the gradient micro-inertia $\bar{\eta} \|\nabla u_{,t}\|^2$ to the classical terms $\rho \|u_{,t}\|^2 + \eta \|P_{,t}\|^2$. We

recall that the strain energy density for this model in the isotropic case takes the form

$$\begin{aligned}
 W = & \underbrace{\mu_e \|\operatorname{sym}(\nabla u - P)\|^2 + \frac{\lambda_e}{2} (\operatorname{tr}(\nabla u - P))^2}_{\text{isotropic elastic-energy}} + \underbrace{\mu_c \|\operatorname{skew}(\nabla u - P)\|^2}_{\text{rotational elastic coupling}} \\
 & + \underbrace{\mu_{\text{micro}} \|\operatorname{sym} P\|^2 + \frac{\lambda_{\text{micro}}}{2} (\operatorname{tr} P)^2}_{\text{micro-self-energy}} + \underbrace{\frac{\mu_e L_c^2}{2} \|\nabla P\|^2}_{\text{isotropic curvature}},
 \end{aligned} \tag{19}$$

The dynamical equilibrium equations are

$$\rho u_{,tt} - \operatorname{Div}[\mathcal{I}] = \operatorname{Div}[\tilde{\sigma}], \quad \eta P_{,tt} = \tilde{\sigma} - s - M, \tag{20}$$

where

$$\begin{aligned}
 \mathcal{I} &= \bar{\eta}_1 \operatorname{dev} \operatorname{sym} \nabla u_{,tt} + \bar{\eta}_2 \operatorname{skew} \nabla u_{,tt} + \frac{1}{3} \bar{\eta}_3 \operatorname{tr}(\nabla u_{,tt}), \\
 \tilde{\sigma} &= 2\mu_e \operatorname{sym}(\nabla u - P) + \lambda_e \operatorname{tr}(\nabla u - P) \mathbb{1} + 2\mu_c \operatorname{skew}(\nabla u - P), \\
 s &= 2\mu_{\text{micro}} \operatorname{sym} P + \lambda_{\text{micro}} \operatorname{tr}(P) \mathbb{1}, \\
 M &= -\mu_e L_c^2 \underbrace{\operatorname{Div} \nabla P}_{=\Delta P}.
 \end{aligned} \tag{21}$$

Recalling the results of Madeo et al. (2014), we remark that when the gradient micro-inertia is vanishing ($\bar{\eta}_1 = \bar{\eta}_2 = \bar{\eta}_3 = 0$), the Mindlin-Eringen model does not allow the description of band-gaps (see Fig. 9), due to the presence of a straight acoustic wave. On the other hand, when switching on the parameters $\bar{\eta}_2$ and $\bar{\eta}_3$, some optic branches are flattened, so that a band-gap can be created suitably choosing the values of the characteristic cutoff frequencies (see Fig. 10). The analogous case for the relaxed micromorphic model (Fig. 5) allowed instead for the description of 2 band-gaps.

In conclusion, the classical Mindlin-Eringen model is suitable to account for dispersion, and for non-localities in metamaterials, it shows its limits for the description of band-gaps in metamaterials. In fact, the presence of a gradient micro-inertia is needed to create a single band-gap, while the relaxed micromorphic model is able to account for 2 band-gaps in the analogous case (see Fig. 6). This means that the relaxed micromorphic model is most suitable to model mechanical metamaterials exhibiting band-gaps also when they are manufactured in such a way that the gradient micro-inertia becomes non-negligible. Indeed, the description of the second band-gap occurring at higher frequencies becomes possible for the relaxed micromorphic model, but not for the Mindlin-Eringen one when considering suitable additional micro-inertia terms.

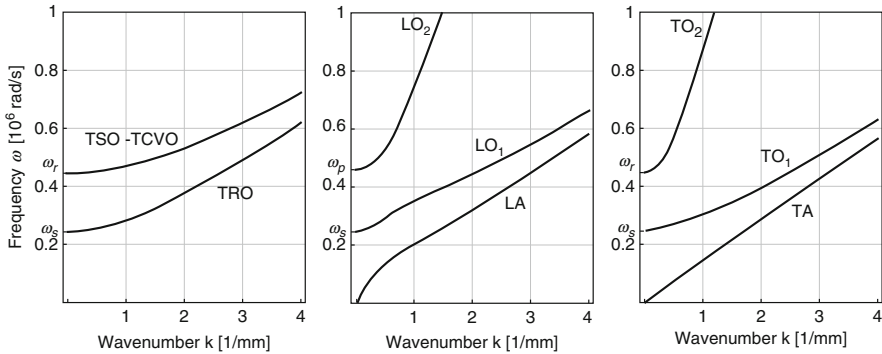


Fig. 9 Dispersion relations $\omega = \omega(k)$ of the **standard Mindlin-Eringen micromorphic model** for the uncoupled (*left*), longitudinal (*center*), and transverse (*right*) waves with nonvanishing free micro-inertia $\eta \neq 0$ and vanishing gradient micro-inertia $\bar{\eta} = 0$

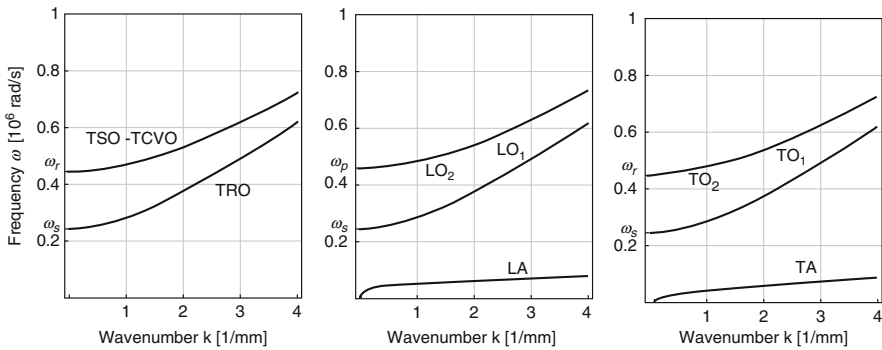


Fig. 10 Dispersion relations $\omega = \omega(k)$ of the **standard Mindlin-Eringen micromorphic model** for the uncoupled (*left*), longitudinal (*center*), and transverse (*right*) waves with nonvanishing free micro-inertia $\eta \neq 0$ and nonvanishing gradient micro-inertia $\bar{\eta} \neq 0$

The Micromorphic Model with Curvature $\|\text{Div}P\|^2 + \|\text{Curl}P\|^2$

The classical Mindlin-Eringen model can be reformulated by introducing the energy (see Madeo et al. 2016c):

$$\begin{aligned}
 W = & \underbrace{\mu_e \|\text{sym}(\nabla u - P)\|^2 + \frac{\lambda_e}{2} (\text{tr}(\nabla u - P))^2}_{\text{isotropic elastic-energy}} + \underbrace{\mu_c \|\text{skew}(\nabla u - P)\|^2}_{\text{rotational elastic coupling}} \\
 & + \underbrace{\mu_{\text{micro}} \|\text{sym} P\|^2 + \frac{\lambda_{\text{micro}}}{2} (\text{tr} P)^2}_{\text{micro-self-energy}} + \underbrace{\frac{\mu_e L_c^2}{2} (\|\text{Div} P\|^2 + \|\text{Curl} P\|^2)}_{\text{augmented isotropic curvature}},
 \end{aligned}
 \tag{22}$$

in which the gradient of the micro-distortion tensor is decomposed in its Curl and Div part. The dynamical equilibrium equations are

$$\rho u_{,tt} - \underbrace{\text{Div}[\mathcal{I}]}_{\text{new augmented term}} = \text{Div}[\tilde{\sigma}], \quad \eta P_{,tt} = \tilde{\sigma} - s - M, \quad (23)$$

where

$$\begin{aligned} \mathcal{I} &= \bar{\eta}_1 \text{dev sym } \nabla u_{,tt} + \bar{\eta}_2 \text{skew } \nabla u_{,tt} + \frac{1}{3} \bar{\eta}_3 \text{tr}(\nabla u_{,tt}), \\ \tilde{\sigma} &= 2 \mu_e \text{sym}(\nabla u - P) + \lambda_e \text{tr}(\nabla u - P) \mathbb{1} + 2 \mu_c \text{skew}(\nabla u - P), \quad (24) \\ s &= 2 \mu_{\text{micro}} \text{sym} P + \lambda_{\text{micro}} \text{tr}(P) \mathbb{1}, \\ M &= -\mu_e L_c^2 \underbrace{(\nabla(\text{Div} P) - \text{Curl Curl} P)}_{\text{Div} \nabla P = \Delta P}. \end{aligned}$$

Simply rearranging the terms appearing in the definition of the tensor M , it can be remarked that the structure of the equation is equivalent to the one obtained in the standard micromorphic model with curvature $\frac{1}{2} \|\nabla P\|^2$; see Eq. (20) in section “[The Standard Mindlin-Eringen Model](#).” The results obtained for such case are thus completely superposable to those presented before for the classical Mindlin-Eringen micromorphic model (see Figs. 9 and 10).

The Micromorphic Model with Curvature $\|\text{Div} P\|^2$

We have shown in the previous sections that the relaxed micromorphic model with curvature $\|\text{Curl} P\|^2$ is an enriched model which allows to simultaneously account for non-localities and multiple band-gaps in mechanical metamaterials. We also showed that the full Mindlin-Eringen model with curvature $\|\nabla P\|^2$ can be equivalently reformulated with a curvature term of the type $\|\text{Curl} P\|^2 + \|\text{Div} P\|^2$. We concluded that the relaxed micromorphic model is well-adapted for the description of band-gap metamaterials, while the classical Mindlin-Eringen model, if equally suitable for describing dispersion and non-locality, is not performant for the modeling of band-gap metamaterials to the same extent. In this section we explore another possible micromorphic model with curvature term $\|\text{Div} P\|^2$ (see Madeo et al. 2016c) which thus features an energy of the type

$$\begin{aligned} W &= \underbrace{\mu_e \|\text{sym}(\nabla u - P)\|^2 + \frac{\lambda_e}{2} (\text{tr}(\nabla u - P))^2}_{\text{isotropic elastic-energy}} + \underbrace{\mu_c \|\text{skew}(\nabla u - P)\|^2}_{\text{rotational elastic coupling}} \\ &+ \underbrace{\mu_{\text{micro}} \|\text{sym} P\|^2 + \frac{\lambda_{\text{micro}}}{2} (\text{tr} P)^2}_{\text{micro-self-energy}} + \underbrace{\frac{\mu_e L_d^2}{2} \|\text{Div} P\|^2}_{\text{isotropic curvature}}. \end{aligned} \quad (25)$$

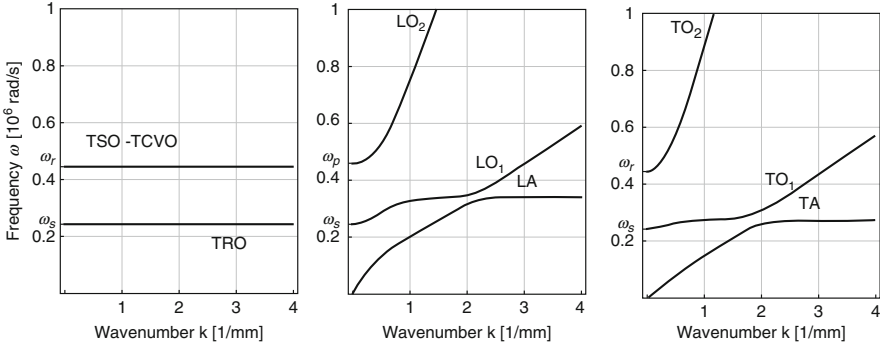


Fig. 11 Dispersion relations $\omega = \omega(k)$ of the **relaxed micromorphic model with curvature** $\|\text{Div} P\|^2$ for the uncoupled (*left*), longitudinal (*center*), and transverse (*right*) waves with nonvanishing free micro-inertia $\eta \neq 0$ and vanishing gradient micro-inertia $\bar{\eta} = 0$

The dynamical equilibrium equations are

$$\rho u_{,tt} - \text{Div}[\mathcal{I}] = \text{Div}[\tilde{\sigma}], \quad \eta P_{,tt} = \tilde{\sigma} - s - M, \quad (26)$$

where

$$\begin{aligned} \mathcal{I} &= \bar{\eta}_1 \text{dev sym } \nabla u_{,tt} + \bar{\eta}_2 \text{skew } \nabla u_{,tt} + \frac{1}{3} \bar{\eta}_3 \text{tr}(\nabla u_{,tt}), \\ \tilde{\sigma} &= 2 \mu_e \text{sym}(\nabla u - P) + \lambda_e \text{tr}(\nabla u - P) \mathbb{1} + 2 \mu_c \text{skew}(\nabla u - P), \quad (27) \\ s &= 2 \mu_{\text{micro}} \text{sym} P + \lambda_{\text{micro}} \text{tr}(P) \mathbb{1}, \\ M &= -\mu_e L_c^2 \nabla(\text{Div} P). \end{aligned}$$

We start by presenting in Fig. 11 the dispersion relations obtained with a vanishing gradient micro-inertia (Fig. 11). Also in the case of the micromorphic model with only $\|\text{Div} P\|^2$, when considering a vanishing gradient micro-inertia, there always exist waves which propagate inside the considered medium independently of the value of the frequency, and the uncoupled waves assume a peculiar resonant behavior in which the frequency is independent of the wavenumber k (local resonances for some modes). The characteristic behavior observed for the longitudinal and transverse waves shows a significant coupling between the LO_1 and LA as well between as the TO_1 and TA waves, respectively. Such strong coupling does not allow for the presence of complete band-gaps. Moreover, since the non-localities introduced by the curvature $\|\text{Div} P\|^2$ are much weaker than those introduced by the term $\|\text{Curl} P\|^2$, the behavior of the dispersion curves for the uncoupled waves is the same as that of a completely local model as the internal

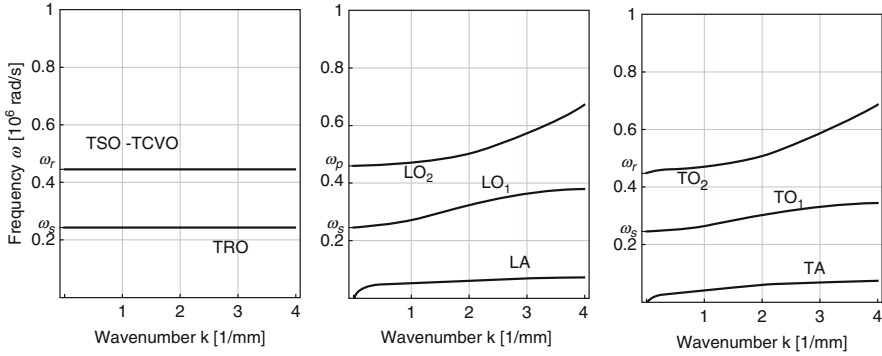


Fig. 12 Dispersion relations $\omega = \omega(k)$ of the **relaxed micromorphic model with curvature** $\|\text{Div} P\|^2$ for the uncoupled (*left*), longitudinal (*center*), and transverse (*right*) waves with nonvanishing free micro-inertia $\eta \neq 0$ and nonvanishing gradient micro-inertia $\bar{\eta} \neq 0$

variable one (see the first of Fig. 7) and significantly deviates from the behavior observed for the relaxed micromorphic model (see the first of Fig. 5).

On the other hand, when switching on the gradient inertia (Fig. 12), a behavior analogous to the relaxed micromorphic model appears for the longitudinal and transverse waves, while the uncoupled waves definitely remain those of a local model. Thus, if the micromorphic model with curvature $\|\text{Div} P\|^2$ can describe band-gaps when adding the gradient micro-inertia terms, it loses almost any capability of accounting for non-local effects. Indeed, the effect of the addition of the gradient micro-inertia is that of flattening the longitudinal and transverse waves which is equivalent to say that these waves become more “local.” The main difference between Figs. 12 and 6 can be found in the uncoupled waves: if in the relaxed micromorphic model they are able to account for non-negligible non-local effects, this is not the case for the micromorphic model with curvature $\|\text{Div} P\|^2$.

We can conclude that, independently of the type of micro-inertia which is considered, the micromorphic model with curvature $\|\text{Div} P\|^2$ is not able to simultaneously account for band-gaps and non-localities, in strong contrast to what happens for the relaxed micromorphic case.

Conclusions

It has been shown in this contribution that the dispersive behaviors usually encountered in engineering metamaterials can be suitably accounted for by using enriched continuum models of the micromorphic type. The enriched kinematics that is peculiar of such enriched models allows for the introduction of extra degrees of freedom with respect to classical Cauchy continua, thus accounting for the description of independent motions of the microstructure inside the unit cells. Nevertheless, if all are suitable for modeling dispersive phenomena in microstructured materials, not all

such models are equally adapted for describing the onset of band-gaps in mechanical metamaterials. The models proposed in this contribution differ one from the other basically for the expression of the curvature terms in the strain energy density, i.e., for the type of higher-order space derivatives of the micro-distortion tensor P which are accounted for.

After presenting in detail the dispersion curves associated to all the introduced isotropic, linear-elastic models, it is concluded that the relaxed micromorphic model is the most suitable enriched model which is simultaneously able to account for both band-gaps and non-localities in mechanical metamaterials.

The interest of introducing complex micro-inertia terms (gradient micro-inertia) accounting for a coupling of the motion of the microstructure with the deformation of the unit cell at the higher scale is also discussed. It is shown that the effect of adding such gradient micro-inertia is that of flattening the longitudinal and transverse dispersion curves, so that the final effect is that of lowering the speed of propagation of such waves at the macroscopic level. This can be equivalently interpreted as a “loss of non-locality” for the longitudinal and transverse waves which is substantially related to the presence of the gradient micro-inertia. As far as the uncoupled waves are concerned, their behavior is not affected by the addition of such new micro-inertia term, and the peculiar behavior of these waves for the relaxed micromorphic model makes evident the non-locality that such model is able to provide concomitantly to the presence of complete band-gaps.

The main interest of using continuum theories of the micromorphic type for the description of the behavior of materials with complex microstructures can be found in the fact that they feature the introduction of few parameters which are, in an averaged sense, reminiscent of the presence of the underlying microstructure. If, on the one hand, this fact provides a drastic modeling simplification which is optimal to proceed toward (meta-)structural design, some drawbacks can be reported which are mainly related to the difficulty of directly relating the introduced macroscopic parameters to the specific characteristics of the microstructure (topology, microstructural mechanical properties, etc.). The aforementioned difficulty of explicitly relating macro-parameters to micro-properties is often seen as a limitation for the effective application of enriched continuum models. As a matter of fact, it is the authors’ belief that such models are a necessary step if one wants to proceed toward the engineering design of metastructures, i.e., structures which are made up of metamaterials as building blocks. Of course, the proposed model will introduce a certain degree of simplification, but it is exactly this simplicity that makes possible to envision the next step which is that of proceeding toward the design of complex (meta-) structures made of metamaterials.

To be more precise, the relaxed micromorphic model proposed here is able to describe the onset of the first (and sometimes the second) band-gap which occurs at lower frequencies. In order to catch more complex behaviors, the kinematics and the constitutive relations of the proposed model should be further enriched in a way that is not yet completely clear. Nevertheless, this fact has not to be seen as a limitation since it is possible to use the unorthodox dynamical behavior of some metamaterials exhibiting band-gaps to fit, by inverse approach, the parameters of

the relaxed micromorphic model following what has been done, e.g., in Madeo et al. (2016a). This fitting, when successfully concluded for some specific metamaterials, will allow the setting up of the design of metastructures by means of tools which are familiar to engineers, such as finite element codes. Of course, as classical Cauchy models show their limits for the description of the dynamical behavior of metamaterials, even at low frequencies, the relaxed micromorphic model will show its limit for higher frequencies, yet remaining accurate enough for accounting for some important macroscopic manifestations of microstructure.

Acknowledgements Angela Madeo thanks the Institut Universitaire de France (IUF) for financial support, INSA-Lyon for the funding of the BQR 2016 “Caractérisation mécanique inverse des métamatériaux: modélisation, identification expérimentale des paramètres et évolutions possibles,” as well as the CNRS-INSIS for the funding of the PEPS project.

References

- J.D. Achenbach, *Wave Propagation in Elastic Solids* (North-Holland Publishing Company, Amsterdam, 1973)
- M.N. Armenise, C.E. Campanella, C. Ciminelli, F. Dell’Olio, V.M.N. Passaro, Phononic and photonic band gap structures: modelling and applications. *Phys. Procedia* **3**(1), 357–364 (2010)
- G. Barbagallo, M.V. d’Agostino, R. Abreu, I.-D. Ghiba, A. Madeo, P. Neff, Transparent anisotropy for the relaxed micromorphic model: macroscopic consistency conditions and long wave length asymptotics *International Journal of Solids and Structures*, <https://doi.org/10.1016/j.ijsolstr.2017.01.030>
- M.V. d’Agostino, G. Barbagallo, I.-D. Ghiba, R. Abreu, A. Madeo, P. Neff, A panorama of dispersion curves for the isotropic weighted relaxed micromorphic model *ZAMM*, DOI: <https://doi.org/10.1002/zamm.201600227>
- A.C. Eringen, Mechanics of micromorphic materials, in *Applied Mechanics* (Springer, Berlin/Heidelberg, 1966), pp. 131–138
- A.C. Eringen, *Microcontinuum Field Theories* (Springer, New York, 1999)
- I.-D. Ghiba, P. Neff, A. Madeo, L. Placidi, G. Rosi, The relaxed linear micromorphic continuum: existence, uniqueness and continuous dependence in dynamics. *Math. Mech. Solids* **20**(10), 1171–1197 (2014)
- A. Madeo, P. Neff, I.-D. Ghiba, L. Placidi, G. Rosi, Band gaps in the relaxed linear micromorphic continuum. *Zeitschrift für Angewandte Mathematik und Mechanik* **95**(9), 880–887 (2014)
- A. Madeo, P. Neff, I.-D. Ghiba, L. Placidi, G. Rosi, Wave propagation in relaxed micromorphic continua: modeling metamaterials with frequency band-gaps. *Contin. Mech. Thermodyn.* **27**(4–5), 551–570 (2015)
- A. Madeo, G. Barbagallo, M.V. d’Agostino, L. Placidi, P. Neff, First evidence of non-locality in real band-gap metamaterials: determining parameters in the relaxed micromorphic model. *Proc. R. Soc. A Math. Phys. Eng. Sci.* <https://doi.org/10.1098/rspa.2016.0169> **472**(2190)
- A. Madeo, P. Neff, E.C. Aifantis, G. Barbagallo, M.V. d’Agostino, On the role of micro-inertia in enriched continuum mechanics *Proceedings of the Royal Society A*, <https://doi.org/10.1098/rspa.2016.0722>
- A. Madeo, P. Neff, M.V. d’Agostino, G. Barbagallo, Complete band gaps including non-local effects occur only in the relaxed micromorphic model. *Comptes Rendus Mécanique*, **344**(11–12), 784–796
- A. Madeo, P. Neff, I.-D. Ghiba, G. Rosi, Reflection and transmission of elastic waves in non-local band-gap metamaterials: a comprehensive study via the relaxed micromorphic model. *J. Mech. Phys. Solids* **95**, 441–479 (2016d)

- W. Man, M. Florescu, K. Matsuyama, P. Yadak, G. Nahal, S. Hashemizad, E. Williamson, P. Steinhardt, S. Torquato, P. Chaikin, Photonic band gap in isotropic hyperuniform disordered solids with low dielectric contrast. *Opt. Express* **21**(17), 19972–19981 (2013)
- R.D. Mindlin, Microstructure in linear elasticity. Technical report, Office of Naval Research, 1963
- R.D. Mindlin, Micro-structure in linear elasticity. *Arch. Ration. Mech. Anal.* **16**(1), 51–78 (1964)
- P. Neff, I.-D. Ghiba, M. Lazar, A. Madeo, The relaxed linear micromorphic continuum: well-posedness of the static problem and relations to the gauge theory of dislocations. *Q. J. Mech. Appl. Math.* **68**(1), 53–84 (2014a)
- P. Neff, I.-D. Ghiba, A. Madeo, L. Placidi, G. Rosi, A unifying perspective: the relaxed linear micromorphic continuum. *Contin. Mech. Thermodyn.* **26**(5), 639–681 (2014b)
- K. Pham, V.G. Kouznetsova, M.G.D. Geers, Transient computational homogenization for heterogeneous materials under dynamic excitation. *J. Mech. Phys. Solids* **61**(11), 2125–2146 (2013)
- A. Sridhar, V.G. Kouznetsova, M.G.D. Geers, Homogenization of locally resonant acoustic metamaterials towards an emergent enriched continuum. *Comput. Mech.* **57**(3), 423–435 (2016)
- G. Stefano, M.S. Greene, W.K. Liu, Characterization of heterogeneous solids via wave methods in computational microelasticity. *J. Mech. Phys. Solids* **58**(5), 959–974 (2011)
- W. Steurer, D. Sutter-Widmer, Photonic and phononic quasicrystals. *J. Phys. D Appl. Phys.* **40**(13), 229–247 (2007)

Part III
**Mathematical Methods in Nonlocal
Continuum Mechanics**



Implicit Nonlocality in the Framework of Viscoplasticity

21

Wojciech Sumelka and Tomasz Łodygowski

Contents

Introduction	744
Relaxation Time: Implicit Length Scale Parameter	746
The Thermo-viscoplasticity Model	747
Basic Definitions	747
Kinematics	749
Constitutive Postulates	754
Constitutive Relations	754
The Microdamage Tensor	756
Definition of Material Functions for Adiabatic Process	760
Evolution Equations for Internal State Variables	760
Material Functions	760
Numerical Examples: Modeling of Spalling Phenomena	765
Introductory Remarks	765
Initial Boundary Value Problem	765
Material Parameter Identification for HSLA-65 Steel	770
Computer Implementation in Abaqus/Explicit	770
Spall Fracture Phenomenon Modeling	771
Conclusions	774
References	776

Abstract

The considerations are addressed to the notion of implicit nonlocality in mechanical models. The term *implicit* means that there is no direct measure of nonlocal action in a model (like classical or fractional gradients, etc. in explicit nonlocal models), but some phenomenological material parameters can be interpreted as

W. Sumelka (✉) · T. Łodygowski
Institute of Structural Engineering, Poznan University of Technology, Poznan, Poland
e-mail: wojciech.sumelka@put.poznan.pl

one that maps some experimentally observed phenomena responsible for the scale effects.

The overall discussion is conducted in the framework of the Perzyna Theory of Viscoplasticity where the role of the implicit length scale parameter plays the relaxation time of the mechanical disturbance. In this sense, in the viscoplastic range of the material behavior, the deformation at each material point contributes to the finite surrounding. The important consequence is that the solution of the IBVP described by Perzyna's theory is unique – the relaxation time is the regularizing parameter.

Keywords

Implicit nonlocality · Viscoplasticity · Anisotropic damage

Introduction

Fifty years after publishing the fundamental paper on the theory of thermo-viscoplasticity (TTV) by Perzyna (1963), the concept of material overstress function is still vivid nowadays in mechanics and enables new findings (cf. Sumelka 2014; Glema et al. 2014; Sumelka and Nowak 2016). One can say that the viscoplasticity concept is one of the most fruitful ideas in modeling plastic (irreversible) processes. It is obvious that TTV formulation evolved during the years to obtain the mature form in the early 1990s (Perzyna 2005) – the one strongly influenced by the works of J.E. Marsden group (cf. Marsden and Hughes 1983; Abraham et al. 1988) – where both detailed experimental observations and robust mathematical modeling have aimed to obtain the unique solution of the posed thermomechanical problem with clear physical interpretation.

The last mentioned aspect of uniqueness of initial boundary value problem (IBVP) formulated in the framework of TTV plays the fundamental role in the following discussion. This crucial aspect involves the concept of the relaxation time – the parameter which controls viscoplasticity effects in TTV and maps implicitly the physical length scale. On the other hand, the main drawback of TTV is complexity of the formulation that requires the necessity of dynamically solving what is computationally time-consuming (Łodygowski 1996). Herein it should be pointed out that TTV and its inherent properties strongly differ from the classical plasticity with application of yield strength limit in a rate form (Heeres et al. 2002) – in this sense the term ‘viscoplasticity’ is used misleadingly in the literature (thus the reader must carefully analyze the definition of plastic strain in a specific formulation to judge whether it is consistent with original TTV).

TTV was extensively verified and validated in the literature. Let us mention herein some crucial aspects of TTV discussed by the authors in a series of papers: (i) mathematically well-posedness was presented in Łodygowski et al. (1994), Łodygowski (1996) and Glema et al. (1997); (ii) solution existence and uniqueness of softening problem were subjects of Perzyna (2005); (iii) viscosity

defined by material parameter was analyzed in Glema (2000) and Glema et al. (2003); (iv) propagation of mechanical and thermal waves was clarified in Glema and Łodygowski (2002) and Glema (2004); (v) dispersive material character was presented in Glema (2000); (vi) diverse way of energy dissipation was discussed in Glema et al. (2003); (vii) smooth and non-smooth distributions within damage and failure were concluded in Glema et al. (2009); (viii) damage anisotropy was included into TTV in Perzyna (2008) and Glema et al. (2009) and developed in Sumelka (2009); (ix) transition from ductile to brittle type of damage and role of covariance in damage mechanics was analyzed in Łodygowski and Sumelka (2015) and Sumelka (2013); (x) or finally the generalization of classical TTV by fractional calculus (Podlubny 1999; Kilbas et al. 2006) application to obtain *fractional viscoplasticity* was recently proposed in Sumelka (2014) and since then has been under continuous development, e.g., Sumelka and Nowak (2017), Sun and Shen (2017) and Xiao et al. (2017). All these results prove that TTV is a reliable tool for modeling the varied types of materials including geometerials, concrete materials, and especially metallic-like materials for a broad range of strains, strain rates, and temperatures.

It should be stated that the explicit nonlocal mechanical models like those proposed by Mindlin and Tiersten (1962), Kröner (1963), Toupin (1963, 1964), Green and Rivlin (1964), Mindlin (1964; 1965) and Mindlin and Eshel (1968), Yang et al. (2002), and Park and Gao (2008) and the nonlocal continuum mechanics initiated by Eringen and coworkers (Eringen 1972a,b, 1983), generalized continuum formulations (Polyzos and Fotiadis 2012; Tarasov 2014), or the one by de Borst and Pamin (1996), Fleck and Hutchinson (1997), Aifantis (1999), and Voyiadjis and Abu Al-Rub (2005) are more robust than implicit TTV (other implicit models can be found in, e.g., Voyiadjis and Abed 2005, 2006; Voyiadjis and Kattan 2007). This means that for the explicit nonlocal formulations, the advantages of inclusion of the length scale parameter are more pronounced and physical interpretation is more straightforward. Nonetheless, the serious drawback of explicit models is the high number of material parameters; recall that Mindlin theory contains 1764 coefficients in total (903 independent, Morán 2016) and these concepts make the numerical implementation more involved and computation time can be significantly higher due to the additional variables (e.g., higher order stresses).

The remaining part of this paper is organized as follows.

In section “[Relaxation Time: Implicit Length Scale Parameter](#),” the physical interpretation of the relaxation time of a mechanical perturbation as a length scale parameter is provided.

Section “[The Thermo-viscoplasticity Model](#)” describes the TTV accounting for anisotropic damage nucleation and growth (Sumelka 2009).

In section “[Definition of Material Functions for Adiabatic Process](#)” the material function for an adiabatic processes are identified.

Finally, in section “[Numerical Examples: Modeling of Spalling Phenomena](#),” the numerical results of spalling phenomena including the effects of anisotropic damage nucleation and growth are presented.

Section “[Conclusions](#)” concludes the document.

Relaxation Time: Implicit Length Scale Parameter

Herein the original concepts by Perzyna are recalled (2005, 2010, 2012).

The attention is focused on metallic materials. It is important that basic results come from the single crystal deformation analysis (micro scale of observation) and afterward they are generalized for the polycrystalline solids (meso-macro scale of observation), the one captured by the continuum description.

From the theory of crystal dislocations, the inelastic shear strain rate can be expressed as

$$\dot{\epsilon}^p = \alpha b v, \quad (1)$$

where α denotes the mean density of mobile dislocations, b is the Burgers vector (the displacement per dislocation line), and v is the mean dislocation velocity (Asaro 1983). Next, based on the experimental observations that thermally activated and phonon damping mechanisms are most pronounced, the mean dislocation velocity can be expressed by

$$v = \frac{AL^{-1}}{t_S + t_B}, \quad (2)$$

where AL^{-1} is the average distance of dislocation movement after each thermal activation, t_S is the time a dislocation has spent at the obstacle, and t_B denotes the time of traveling between the barriers. Applying Eq. (2) in Eq. (1), it can be shown that (Teodosiu and Sidoroff 1976)

$$\begin{aligned} \dot{\epsilon}^p &= \frac{1}{T_{mT}} \left\langle \exp \left\{ \frac{U[(\tau - \tau_\mu)Lb]}{k\vartheta} \right\} + \frac{BAL^{-1}v}{(\tau - \tau_B)b} \right\rangle^{-1} \\ &= \frac{1}{T} \left\langle \Phi \left[\frac{\tau}{\tau_Y} - 1 \right] \right\rangle \text{sgn}\tau, \end{aligned} \quad (3)$$

where T_{mT} is the relaxation time for the thermally activated mechanism, U is the activation energy (Gibbs free energy), τ is the applied stress, τ_μ is the athermal stress, L is the distance between obstacle dislocation, k is the Boltzmann constant, B is called the dislocation drag coefficient, ϑ is the absolute temperature, v is the frequency of dislocation vibration, τ_B denotes the stress needed to overcome the forest dislocation barriers to the dislocation motion (the back stress), $\langle \cdot \rangle$ denotes the Macaulay bracket, T is the relaxation time, Φ is the empirical overstress function, and τ_Y is a static yield stress. From the point of view of herein discussion, it is crucial that

$$T_{mT} = \frac{1}{\alpha b AL^{-1}v}. \quad (4)$$

Therefore, the relaxation time for the thermally activated mechanism is a function of the average distance of dislocation movement, thus including the information on the characteristic size of plastic deformation, identified with crystal (micro) length scale characteristic for specific material. For completeness it should be stated that the relaxation time for the phonon damping mechanism (T_{mD}) is

$$T_{mD} = \frac{B}{\alpha b^2 \tau_B} = T_{mT} \frac{BAL^{-1}v}{b\tau_B}. \quad (5)$$

The final postulate is that for a polycrystalline material (continuum level), the meso-macro viscoplastic strain is described by the analogous formula to Eq. (1). As will be discussed in the following section, the macroscopic relaxation time T_m is then

$$T_m = \frac{\ell}{\beta c}, \quad (6)$$

where ℓ is a macroscopic length scale, β is a proportionality factor, and c denotes the velocity of the propagation of the elastic wave. It can be shown Sluys (1992) that for 1D longitudinal wave propagation in the elasto-viscoplastic material

$$T_m = \frac{\ell E}{2\sigma_0 c}, \quad (7)$$

where E is the Young modulus and σ_0 is the yield stress; thus, $\beta = \frac{2\sigma_0}{E}$.

The last relation shows clearly that the relaxation time maps the material length scale, thus implicitly introducing length scale to the continuum model. Furthermore, the relaxation time T_m can be viewed not only as a microstructural parameter to be determined from experimental observations but also as a mathematical regularization parameter, as mentioned.

The Thermo-viscoplasticity Model

The assumed kinematics induces the form of stresses (Dłuzewski 1996); hence, first the kinematics of the body is described. Next, following the balance principles and necessary constitutive axioms, the constitutive model for a thermomechanical process is obtained. It should be pointed out that the description follows the concepts presented in Sumelka (2009).

Basic Definitions

The following fundamental definitions hold in continuum mechanics (Truesdell and Noll 1965; Rymarz 1993; Perzyna 1978; Ostrowska-Maciejewska 1994).

Definition 1. The **material continuum** is a three-dimensional differentiable manifold \mathcal{M} :

$$\mathcal{M} = (\mathbb{R}^3, S)$$

where \mathbb{R}^3 is a three-dimensional continuum and S denotes the structure of manifold.

Definition 2. The material **body** \mathcal{B} is a subset of material continuum \mathcal{M} :

$$\mathcal{B} \subset \mathcal{M}$$

and is characterized by the following structure S :

1. The set of material bodies is a material body:

$$\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \dots \cup \mathcal{B} \subset \mathcal{M},$$

2. Material body \mathcal{B} or its arbitrary part can be uniquely mapped in the finite domain D of the Euclidean point space E^3 :

$$\boldsymbol{\kappa} : \mathcal{B} \rightarrow D,$$

where $\boldsymbol{\kappa} \subset E^3$. The map $\boldsymbol{\kappa}$ is called the **configuration** of the body and $\boldsymbol{\kappa} \in K$, where K is a set of a homeomorphic maps.

3. The smallest part of the material body is a **material point** X . The map $\boldsymbol{\kappa}$ assigns to every material point, X , its geometrical point in D

$$\mathbf{X} = \boldsymbol{\kappa}(X), \quad X = \boldsymbol{\kappa}^{-1}(\mathbf{X}), \quad \mathbf{X} \in E^3, \quad X \in \mathcal{B}.$$

4. For every pair of configurations $\boldsymbol{\kappa}, \boldsymbol{\gamma}$, there exist a homeomorphic map:

$$\boldsymbol{\phi}_{\boldsymbol{\kappa}\boldsymbol{\gamma}} : \boldsymbol{\gamma} \circ \boldsymbol{\kappa}^{-1}, \quad \boldsymbol{\phi}_{\boldsymbol{\kappa}\boldsymbol{\gamma}} : E^3 \rightarrow E^3.$$

The map $\boldsymbol{\phi}_{\boldsymbol{\kappa}\boldsymbol{\gamma}}$ is called **Deformative configuration** and maps the domain E^3 in E^3 . Those domains are occupied by the body \mathcal{B} in different configurations.

5. In the E^3 space, the **additive Borelian measure** M corresponds to every configuration $\boldsymbol{\kappa}$:

$$M : D \rightarrow \mathbb{R}^1 \quad \text{or} \quad M : \boldsymbol{\kappa}(\mathcal{B}) \rightarrow \mathbb{R}^1.$$

6. The constitutive relations describing material of the body are fulfilled in every material point X of the body \mathcal{B} .

Definition 3. Absolute time is a time which runs at the same rate for all the observers in the universe.

Those definitions state a passage from an abstract mathematical description of the body, \mathcal{B} , to physical one.

Kinematics

Thus, the real material body become the abstract body, modeled as a manifold. The deformation of a real material body is treated in this ‘mathematical world’ as a mapping between manifolds. The observer in an abstract setting can occupy different positions; by analogy to the real observer, however, some of them are most convenient from many point of views. Two of them are most important for description of the material body motion, namely, the Lagrangean (material, referential) and the Eulerian (spatial, current) ones. These descriptions span two manifolds, as mentioned, and will be denoted by \mathcal{B} and \mathcal{S} , respectively (Marsden and Hughes 1983).

Points in \mathcal{B} are denoted by \mathbf{X} while in \mathcal{S} by \mathbf{x} . Coordinate system for \mathcal{B} is denoted by $\{X^A\}$ with base \mathbf{E}_A and for \mathcal{S} by $\{x^a\}$ with base \mathbf{e}_a . Dual bases in those coordinate systems are denoted by \mathbf{E}^A and \mathbf{e}^a , respectively. The tangent spaces in \mathcal{B} and \mathcal{S} are written as $T_{\mathbf{X}}\mathcal{B} = \{\mathbf{X}\} \times V^3$ and $T_{\mathbf{x}}\mathcal{S} = \{\mathbf{x}\} \times V^3$. They are understood as Euclidean vector space V^3 , regarded as vectors emanating from points \mathbf{X} and \mathbf{x} , respectively (Marsden and Hughes 1983).

For measuring purposes in the abstract space, the Riemannian space on manifolds \mathcal{B} and \mathcal{S} is introduced, i.e., $\{\mathcal{B}, \mathbf{G}\}$ and $\{\mathcal{S}, \mathbf{g}\}$ where metric tensors are defined, as $\mathbf{G} : T\mathcal{B} \rightarrow T^*\mathcal{B}$ and $\mathbf{g} : T\mathcal{S} \rightarrow T^*\mathcal{S}$ where $T\mathcal{B}$ and $T\mathcal{S}$ denote the tangent bundles of \mathcal{B} and \mathcal{S} , respectively, while $T^*\mathcal{B}$ and $T^*\mathcal{S}$ denote their dual tangent bundles. Explicit definitions for metric tensors are then $G_{AB}(\mathbf{X}) = (\mathbf{E}_A, \mathbf{E}_B)_{\mathbf{X}}$ and $g_{ab}(\mathbf{x}) = (\mathbf{e}_a, \mathbf{e}_b)_{\mathbf{x}}$ where $(,)_{\mathbf{X}}$ and $(,)_{\mathbf{x}}$ denote inner product in \mathcal{B} and \mathcal{S} , respectively.

The regular motion of the material body is treated as a series of the immersing of the abstract body \mathcal{B} in the Euclidean point space E^3 (Rymarz 1993), and can be written as

$$\mathbf{x} = \phi(\mathbf{X}, t). \quad (8)$$

Thus, $\phi_t : \mathcal{B} \rightarrow \mathcal{S}$ is a C^1 current configuration of \mathcal{B} in \mathcal{S} , at time t . For the analysis of abstract body deformation, the tangent of ϕ , which defines the two point tensor field \mathbf{F} , called deformation gradient, is considered. \mathbf{F} describes all local deformation properties and is the primary measure of deformation (Perzyna 1978; Holzapfel 2000), thus

$$\mathbf{F}(\mathbf{X}, t) = T\phi = \frac{\partial \phi(\mathbf{X}, t)}{\partial \mathbf{X}}, \quad (9)$$

and using the notion of tangent space

$$\mathbf{F}(\mathbf{X}, t) : T_{\mathbf{X}}\mathcal{B} \rightarrow T_{\mathbf{x}=\phi(\mathbf{X}, t)}\mathcal{S}, \quad (10)$$

so \mathbf{F} is a linear transformation for each $\mathbf{X} \in \mathcal{B}$ and $t \in I \subset \mathbb{R}^1$.

The important properties of \mathbf{F} are a consequence of the assumption that map ϕ is uniquely invertible (smooth homeomorphism) ($\mathbf{X} = \phi^{-1}(\mathbf{x}, t)$); hence, there exists the inverse of deformation gradient:

$$\mathbf{F}^{-1}(\mathbf{x}, t) = \frac{\partial \phi^{-1}(\mathbf{x}, t)}{\partial \mathbf{x}}. \tag{11}$$

Thus, the tensor field \mathbf{F} is non-singular ($\det(\mathbf{F}) \neq 0$), and because of the impenetrability of matter $\det(\mathbf{F}) > 0$. Furthermore \mathbf{F} can be uniquely decomposed into pure stretch and pure rotation, called polar decomposition, namely

$$\mathbf{F} = \mathbf{R}\mathbf{U} = \mathbf{v}\mathbf{R}, \tag{12}$$

where \mathbf{R} is the rotation tensor (unique, proper orthogonal) which measures local orientation and \mathbf{U} and \mathbf{v} define unique, positive definite, symmetric tensors called the right (or material) stretch tensor and the left (or spatial) stretch tensor, respectively (stretch tensors measure the local shape). Using the notion of tangent space, the result is obtained that for each $\mathbf{X} \in \mathcal{B}$, $\mathbf{U}(\mathbf{X}) : T_{\mathbf{X}}\mathcal{B} \rightarrow T_{\mathbf{X}}\mathcal{B}$ and for each $\mathbf{x} \in \mathcal{S}$, $\mathbf{v}(\mathbf{x}) : T_{\mathbf{x}}\mathcal{S} \rightarrow T_{\mathbf{x}}\mathcal{S}$.

From the general class of the Lagrangean and the Eulerian strain measures, defined through one single scale function given by (cf. Hill 1978; Xiao et al. 1998),

$$\mathbf{E} = g(\mathbf{C}) = \sum_{i=1}^3 g(\check{\chi}_i)\mathbf{C}_i,$$

and

$$\mathbf{e} = g(\mathbf{B}) = \sum_{i=1}^3 g(\check{\chi}_i)\mathbf{B}_i,$$

where the scale function $g(\cdot)$ is a smooth increasing function with the normalized property $g(1) = g'(1) - 1 = 0$, $\check{\chi}_i$ is used to denote distinct eigenvalues of the right and left Cauchy-Green tensors \mathbf{C} and \mathbf{B} , respectively, and \mathbf{C}_i and \mathbf{B}_i are the corresponding subordinate eigenprojections; the Green-Lagrange and the Euler-Almansi definitions were accepted. As will be presented, such choice is crucial for the clear meaning of the rate-type constitutive structure.

The Green-Lagrange strain tensor is defined as Perzyna (2005) ($\mathbf{E} : T_{\mathbf{X}}\mathcal{B} \rightarrow T_{\mathbf{X}}\mathcal{B}$):

$$2\mathbf{E} = \mathbf{C} - \mathbf{I}, \tag{13}$$

where \mathbf{E} stands for the Green-Lagrange strain tensor, \mathbf{I} denotes the identity on $T_{\mathbf{X}}\mathcal{B}$, and

$$\mathbf{C} = \mathbf{F}^T \cdot \mathbf{F} = \mathbf{U}^2 = \mathbf{B}^{-1}. \tag{14}$$

On the other side, the Euler-Almansi strain tensor definition is ($\mathbf{e} : T_{\mathbf{x}}\mathcal{S} \rightarrow T_{\mathbf{x}}\mathcal{S}$)

$$2\mathbf{e} = \mathbf{i} - \mathbf{c}, \quad (15)$$

where \mathbf{e} stands for the Euler-Almansi strain tensor and \mathbf{i} denotes the identity on $T_{\mathbf{x}}\mathcal{S}$. One has also

$$\mathbf{c} = \mathbf{b}^{-1} \quad \text{and} \quad \mathbf{b} = \mathbf{F} \cdot \mathbf{F}^T = \mathbf{v}^2, \quad (16)$$

where tensor \mathbf{b} is sometimes referred to as the Finger deformation tensor.

For the purpose of the objective tensor rate definition, the push-forward and the pullback operations are introduced as

$$\phi_*((\cdot)^b) = \mathbf{F}^{-T}(\cdot)^b\mathbf{F}^{-1}, \quad (17)$$

$$\phi^*((\cdot)^b) = \mathbf{F}^T(\cdot)^b\mathbf{F}, \quad (18)$$

respectively, where b indicates that a tensor has all its indices lowered (Marsden and Hughes 1983). In view of the above definitions, the following holds:

$$\mathbf{e}^b = \phi_*(\mathbf{E}^b) = \mathbf{F}^{-T}\mathbf{E}^b\mathbf{F}^{-1}, \quad (19)$$

$$\mathbf{E}^b = \phi^*(\mathbf{e}^b) = \mathbf{F}^T\mathbf{e}^b\mathbf{F}. \quad (20)$$

The generality of possible deformations of a real material body assumes that also in an abstract world it can be arbitrary, called commonly finite. With this respect, the finite elasto-viscoplastic deformation is assumed, manifested by the multiplicative decomposition of the total deformation gradient, namely

$$\mathbf{F}(\mathbf{X}, t) = \mathbf{F}^e(\mathbf{X}, t) \cdot \mathbf{F}^p(\mathbf{X}, t). \quad (21)$$

This decomposition is justified by the micromechanics of single crystal plasticity (Perzyna 1998) and states that the component \mathbf{F}^e is a lattice contribution to \mathbf{F} while \mathbf{F}^p describes the deformation solely due to plastic shearing on crystallographic slip systems.

The inverse of the local elastic deformation \mathbf{F}^{e-1} releases from the stress state in every surrounding ($\mathcal{N}(\mathbf{x}) \subset \phi(\mathcal{B})$) in an actual configuration. The configuration obtained by the linear map \mathbf{F}^{e-1} from actual configuration \mathcal{S} is called unstressed configuration and is denoted by \mathcal{S}' . Thus one can write (see Fig. 1)

$$\mathbf{F}^e : T_{\mathbf{y}}\mathcal{S}' \rightarrow T_{\mathbf{x}}\mathcal{S}, \quad \mathbf{F}^p : T_{\mathbf{x}}\mathcal{B} \rightarrow T_{\mathbf{y}}\mathcal{S}', \quad (22)$$

where the material point in the configuration \mathcal{S}' is characterized by \mathbf{y} .

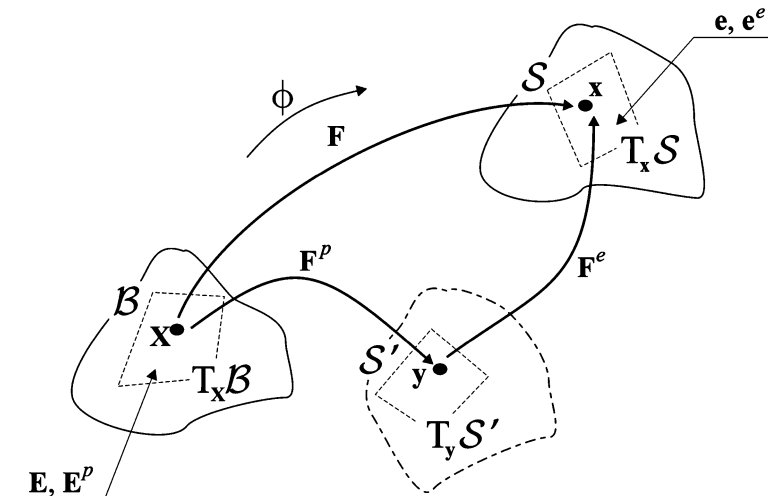


Fig. 1 The interpretation of the multiplicative decomposition of F

Based on the decomposition Eq.(21), the viscoplastic strain tensor $E^p : T_x B \rightarrow T_x B$ can be written as

$$2E^p = C^p - I, \tag{23}$$

where

$$C^p = F^{pT} \cdot F^p = U^{p2} = B^{p-1} \quad \text{and} \quad E^e = E - E^p, \tag{24}$$

whereas the elastic strain tensor $e^e : T_x S \rightarrow T_x S$ is

$$2e^e = i - c^e, \tag{25}$$

where

$$c^e = b^{e-1} \quad \text{and} \quad b^e = F^e \cdot F^{eT} = v^{e2} \quad \text{and} \quad e^p = e - e^e. \tag{26}$$

To define TTV, being the rate-type constitutive structure, the rate of deformation needs to be defined. Starting with the spatial velocity v ,

$$v(x, t) = \dot{x} = \frac{\partial \phi}{\partial t}, \tag{27}$$

the gradient of v , the tensor field (nonsymmetric, second order), called spatial velocity gradient, can be computed (Holzapfel 2000):

$$l(x, t) = \frac{\partial v(x, t)}{\partial x}, \tag{28}$$

where \mathbf{l} stands for spatial velocity gradient. Next, based on the definitions given by Eqs. (9), (21), and (28) the fundamental relation is obtained (Perzyna 2005):

$$\mathbf{l} = \dot{\mathbf{F}} \cdot \mathbf{F}^{-1} = \dot{\mathbf{F}}^e \cdot \mathbf{F}^{e-1} + \mathbf{F}^e \cdot (\dot{\mathbf{F}}^p \cdot \mathbf{F}^{p-1}) \cdot \mathbf{F}^{e-1} = \mathbf{l}^e + \mathbf{l}^p, \quad (29)$$

which introduces the elastic \mathbf{l}^e and plastic \mathbf{l}^p parts of spatial velocity gradient. On the other hand, the additive decomposition of spatial velocity gradient to symmetric and antisymmetric parts generates covariant tensor field \mathbf{d} called rate of deformation tensor and also covariant tensor field \mathbf{w} called spin tensor, with the definitions :

$$\mathbf{l} = \mathbf{d} + \mathbf{w} = \mathbf{d}^e + \mathbf{w}^e + \mathbf{d}^p + \mathbf{w}^p, \quad (30)$$

$$\mathbf{d} = \frac{1}{2}(\mathbf{l} + \mathbf{l}^T), \quad (31)$$

$$\mathbf{w} = \frac{1}{2}(\mathbf{l} - \mathbf{l}^T). \quad (32)$$

To define the objective rate, the Lie-type derivative (assuring diffeomorphisms) is accepted. Its definition for an arbitrary spatial tensor field φ is obtained using the following concept:

- (i) compute the pullback operation of φ – the material field Φ is obtained,
- (ii) take the material time derivative of Φ ,
- (iii) carry out the push-forward operation of the result field from (ii).

The scheme can be summarized as

$$L_v(\varphi) = \phi_* \left(\frac{D}{Dt} \phi^*(\varphi) \right), \quad (33)$$

where L_v stands for Lie derivative.

In view of the above definition, Lie derivative of the Euler-Almansi strain measure results in

$$\mathbf{d}^b = L_v(\mathbf{e}^b). \quad (34)$$

Thus, Lie derivative states a direct relationship between the stretching \mathbf{d} and the Eulerian strain \mathbf{e} . It should be pointed out that Eq. (34) holds for \mathbf{e}^b only; thus, the choice of \mathbf{e}^b to be the strain measure is the most proper from the point of view of physical interpretation of the mathematical model.

Furthermore based on Eq. (34), one can write

$$\mathbf{d}^{e^b} = L_v(\mathbf{e}^{e^b}), \quad \mathbf{d}^{p^b} = L_v(\mathbf{e}^{p^b}). \quad (35)$$

Constitutive Postulates

The constitutive structure must fulfill all basic experimentally proven balance principles. Nonetheless, some constitutive postulates are still needed for completeness. Thus, assuming that conservation of mass, balance of momentum, balance of moment of momentum, and balance of energy and entropy production hold, four constitutive postulates are accepted (Perzyna 1986a, 2005):

- (i) Existence of the free energy function ψ being formally the following scalar function of tensorial argument:

$$\psi = \hat{\psi}(\mathbf{e}, \mathbf{F}, \vartheta; \boldsymbol{\mu}), \quad (36)$$

where $\boldsymbol{\mu}$ denotes a set of internal state variables governing the description of dissipation effects and ϑ represents temperature. It is clear that for nonempty $\boldsymbol{\mu}$ state vector, a dissipation exists in the model; otherwise, the presented model describes thermoelasticity.

- (ii) Axiom of objectivity. The material model should be invariant with respect to diffeomorphism (any superposed motion) – (cf. Frewer 2009) for interesting discussion.
- (iii) The axiom of the entropy production. For every regular process, the constitutive functions should satisfy the second law of thermodynamics.
- (iv) The evolution equation for the internal state variables vector $\boldsymbol{\mu}$ should be of the form

$$\mathbf{L}_v \boldsymbol{\mu} = \hat{\mathbf{m}}(\mathbf{e}, \mathbf{F}, \vartheta, \boldsymbol{\mu}). \quad (37)$$

where evolution function $\hat{\mathbf{m}}$ has to be determined based on the experimental observations. This postulate states the hardest part of modeling; others are nowadays broadly accepted standards.

Constitutive Relations

To obtain the stress and the entropy relations, the reduced dissipation inequality is applied in the form (Marsden and Hughes 1983; Sumelka 2009)

$$\frac{1}{\rho_{\text{Ref}}} \boldsymbol{\tau} : \mathbf{d} - (\eta \dot{\vartheta} + \dot{\psi}) - \frac{1}{\rho \vartheta} \mathbf{q} \cdot \text{grad} \vartheta \geq 0, \quad (38)$$

where ρ denotes current and ρ_{Ref} reference densities, $\boldsymbol{\tau}$ denotes Kirchhoff stress, ψ is the free energy function, ϑ is absolute temperature, η denotes the specific

(per unit mass) entropy, and \mathbf{q} is the heat flux. Using postulate (i), Eq. (38) can be rewritten in the form:

$$\left(\frac{1}{\rho_{\text{Ref}}} \boldsymbol{\tau} - \frac{\partial \hat{\psi}}{\partial \mathbf{e}} \right) : \mathbf{d} - \left(\eta + \frac{\partial \hat{\psi}}{\partial \vartheta} \right) \dot{\vartheta} - \frac{\partial \hat{\psi}}{\partial \boldsymbol{\mu}} \mathbf{L}_v \boldsymbol{\mu} - \frac{1}{\rho \vartheta} \mathbf{q} \cdot \text{grad} \vartheta \geq 0, \quad (39)$$

so because of arbitrariness

$$\boldsymbol{\tau} = \rho_{\text{Ref}} \frac{\partial \hat{\psi}}{\partial \mathbf{e}}, \quad (40)$$

$$\eta = - \frac{\partial \hat{\psi}}{\partial \vartheta}. \quad (41)$$

Thus, Eq. (39) reduces to

$$- \frac{\partial \hat{\psi}}{\partial \boldsymbol{\mu}} \mathbf{L}_v \boldsymbol{\mu} - \frac{1}{\rho \vartheta} \mathbf{q} \cdot \text{grad} \vartheta \geq 0. \quad (42)$$

The specific form of the constitutive relations depends on the assumed internal state vector components. Herein we assume that $\boldsymbol{\mu}$ consists of two variables, namely, (Perzyna 2008; Glema et al. 2009; Sumelka and Łodygowski 2011)

$$\boldsymbol{\mu} = (\epsilon^p, \boldsymbol{\xi}), \quad (43)$$

where ϵ^p is the equivalent plastic deformation $\epsilon^p = \left(\frac{2}{3} \mathbf{d}^p : \mathbf{d}^p \right)^{\frac{1}{2}}$, which describes the dissipation effects generated by viscoplastic deformation and $\boldsymbol{\xi}$ being the microdamage tensor which takes into account the anisotropic microdamage effects.

The assumption Eq. (43), together with the application of Lie derivative to formula Eq. (40), with internal state vector constant (or in other words keeping the history constant – thermoelastic process), gives us the evolution equation for Kirchhoff stress tensor in the form (Duszek–Perzyna and Perzyna 1994)

$$(\mathbf{L}_v \boldsymbol{\tau})^e = \mathcal{L}^e : \mathbf{d}^e - \mathcal{L}^{th} \dot{\vartheta}, \quad (44)$$

where

$$\mathcal{L}^e = \rho_{\text{Ref}} \frac{\partial^2 \hat{\psi}}{\partial \mathbf{e}^2}, \quad (45)$$

$$\mathcal{L}^{th} = - \rho_{\text{Ref}} \frac{\partial^2 \hat{\psi}}{\partial \mathbf{e} \partial \vartheta}, \quad (46)$$

in above \mathcal{L}^e denotes elastic constitutive tensor and \mathcal{L}^{th} is thermal operator. Using the relations

$$(\mathbf{L}_v \boldsymbol{\tau})^e = \dot{\boldsymbol{\tau}} - \boldsymbol{\tau} \cdot \mathbf{d}^e - \mathbf{d}^e \cdot \boldsymbol{\tau}, \quad (47)$$

and

$$\mathbf{d} = \mathbf{d}^e + \mathbf{d}^p, \quad (48)$$

final form of the Kirrchhoff stress rate is

$$\mathbf{L}_v \boldsymbol{\tau} = \mathcal{L}^e : \mathbf{d} - \mathcal{L}^{th} \dot{\vartheta} - (\mathcal{L}^e + \mathbf{g} \boldsymbol{\tau} + \boldsymbol{\tau} \mathbf{g}) : \mathbf{d}^p. \quad (49)$$

On the other side, the energy balance in the form (Perzyna 2005; Sumelka 2009)

$$\rho \vartheta \dot{\eta} = -\text{div} \mathbf{q} - \rho \frac{\partial \hat{\psi}}{\partial \boldsymbol{\mu}} \cdot \mathbf{L}_v \boldsymbol{\mu}, \quad (50)$$

together with the rate of entropy, defined by Eq. (41), defines the evolution equation for temperature, namely,

$$\rho c_p \dot{\vartheta} = -\text{div} \mathbf{q} + \vartheta \frac{\rho}{\rho_{\text{Ref}}} \frac{\partial \boldsymbol{\tau}}{\partial \vartheta} : \mathbf{d} + \chi^* \boldsymbol{\tau} : \mathbf{d}^p + \chi^{**} \mathbf{k} : \mathbf{L}_v \boldsymbol{\xi}, \quad (51)$$

where the specific heat and the irreversibility coefficients χ^* and χ^{**} are determined by

$$c_p = -\vartheta \frac{\partial^2 \hat{\psi}}{\partial \vartheta^2}, \quad (52)$$

$$\chi^* = -\rho \left(\frac{\partial \hat{\psi}}{\partial \epsilon^p} - \vartheta \frac{\partial^2 \hat{\psi}}{\partial \vartheta \partial \epsilon^p} \right) \sqrt{\frac{2}{3}} \frac{1}{\boldsymbol{\tau} : \mathbf{p}}, \quad (53)$$

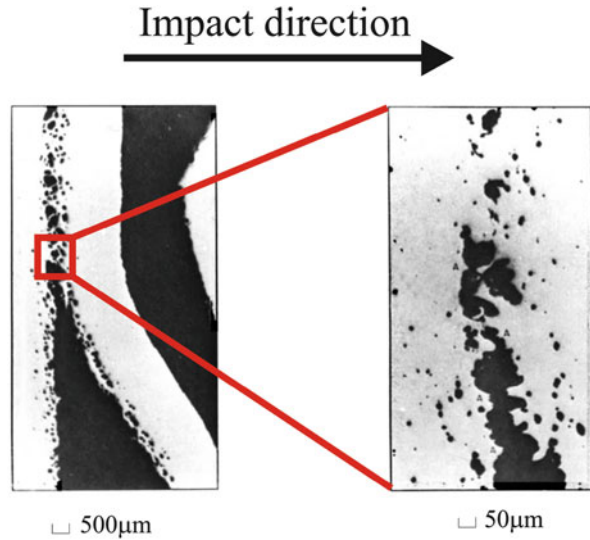
$$\chi^{**} = -\rho \left(\frac{\partial \hat{\psi}}{\partial \boldsymbol{\xi}} - \vartheta \frac{\partial^2 \hat{\psi}}{\partial \vartheta \partial \boldsymbol{\xi}} \right) : \frac{1}{\mathbf{k}},$$

where \mathbf{p} defines the tensor of viscoplastic flow direction.

The Microdamage Tensor

The microdamage tensor is postulated to be the state variable in Eq. (43) and maps the experimentally observed micro-voiding in the zones of the severe plastic

Fig. 2 Cracks anisotropy in 1145 aluminum after flat plate impact experiment (Seaman et al. 1976)



deformation. As an example in Fig. 2, the effects of a flat plate impact experiment in 1145 aluminum are considered (Seaman et al. 1976). It is observed that three stages of damage evolution are crucial: nucleation, growth, and coalescence. Another important fact is that the microdefects have directional geometry – in this experiment they have approximately an ellipsoidal shape. This damage anisotropy is important and should be included in the continuum description (at meso-macro scale), keeping simultaneously the information about level of material porosity.

The microdamage tensor (the state variable at meso-macro scale) Fig. 3 is introduced through the logic presented below, with the most important properties that components of this tensor are proportional to the damage area on the representative volume element (RVE) (its principal values), whereas its Euclidean norm defines porosity.

Considering the RVE, being the abstract point P_i in the model (cf. Fig. 3), one can introduce three ratios:

$$\frac{A_i^p}{A}, \quad (54)$$

where A_i^p is a damaged area and A denotes assumed characteristic area of the RVE. Next, three vectors are obtained based on the introduced ratios – their modules are equal to those ratios and are normal to the RVE's planes. From all possible spatial configuration of the RVE, the one is chosen, in which the resultant module is largest. This resultant is called the *main microdamage vector* and is denoted by $\hat{\xi}^{(m)}$ (Sumelka and Glema 2007), i.e.,

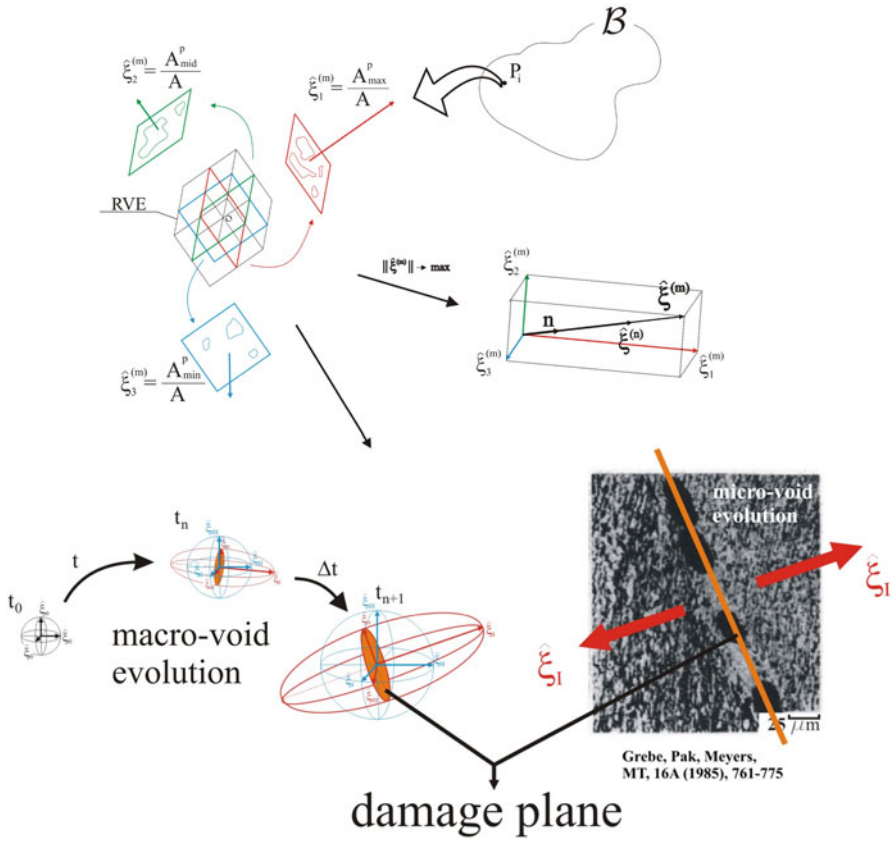


Fig. 3 The concept of microdamage tensor

$$\hat{\xi}^{(m)} = \frac{A_1^p}{A} \hat{\mathbf{e}}_1 + \frac{A_2^p}{A} \hat{\mathbf{e}}_2 + \frac{A_3^p}{A} \hat{\mathbf{e}}_3, \quad (55)$$

where $\hat{\cdot}$ denotes the principal directions of microdamage with $A_1^p \geq A_2^p \geq A_3^p$.

In the following step, based on the main microdamage vector, the *microdamage vector* is built, denoted by $\hat{\xi}^{(n)}$ (Sumelka and Glema 2007):

$$\hat{\xi}^{(n)} = \frac{1}{\|\hat{\xi}^{(m)}\|} \left(\left(\frac{A_1^p}{A} \right)^2 \hat{\mathbf{e}}_1 + \left(\frac{A_2^p}{A} \right)^2 \hat{\mathbf{e}}_2 + \left(\frac{A_3^p}{A} \right)^2 \hat{\mathbf{e}}_3 \right), \quad (56)$$

which states clear connection to microdamage tensor (Sumelka and Glema 2007)

$$\hat{\xi}^{(n)} = \hat{\xi} \mathbf{n}, \quad (57)$$

where

$$\mathbf{n} = \sqrt{3} \left\| \hat{\boldsymbol{\xi}}^{(m)} \right\|^{-1} \left(\hat{\xi}_1^{(m)} \hat{\mathbf{e}}_1 + \hat{\xi}_2^{(m)} \hat{\mathbf{e}}_2 + \hat{\xi}_3^{(m)} \hat{\mathbf{e}}_3 \right), \quad (58)$$

giving in the principal directions of $\boldsymbol{\xi}$ the fundamental relation

$$\hat{\boldsymbol{\xi}} = \frac{\sqrt{3}}{3} \begin{bmatrix} \hat{\xi}_1^{(m)} & 0 & 0 \\ 0 & \hat{\xi}_2^{(m)} & 0 \\ 0 & 0 & \hat{\xi}_3^{(m)} \end{bmatrix}. \quad (59)$$

Therefore, as mentioned, the physical interpretation of the microdamage tensor components is that the diagonal components ξ_{ii} of the microdamage tensor $\boldsymbol{\xi}$, in its principal directions, are proportional to the components of the main microdamage vector $\xi_i^{(m)}$ which defines the ratio of the damaged area to the assumed characteristic area of the RVE, on the plane perpendicular to the i direction. As a consequence, the damage plane is the one perpendicular to the maximal principal value of $\boldsymbol{\xi}$.

Moreover, the Euclidean norm from the microdamage field $\hat{\boldsymbol{\xi}}$ results in additional physical interpretation for microdamage tensor, namely, it defines the scalar quantity called the *volume fraction porosity* or simply *porosity* (Perzyna 2008):

$$\sqrt{\hat{\boldsymbol{\xi}} : \hat{\boldsymbol{\xi}}} = \xi = \frac{V - V_s}{V} = \frac{V_p}{V}, \quad (60)$$

where ξ denotes porosity (scalar damage parameter), V is the volume of a material element, V_s is the volume of the solid constituent of that material element, and V_p denotes void volume:

$$V_p = \frac{\sqrt{3} l}{3} \sqrt{(A_1^p)^2 + (A_2^p)^2 + (A_3^p)^2}. \quad (61)$$

The interpretations of the microdamage tensorial field impose the mathematical bounds for the microdamage evolution, as

$$\xi \in \langle 0, 1 \rangle, \quad \text{and} \quad \hat{\xi}_{ii} \in \langle 0, 1 \rangle. \quad (62)$$

However, the physical bounds are different and are rate dependent (Cochran and Banner 1988; Meyers and Aime 1983), e.g., under extreme loading is of the order $0.09 \div 0.35$ (Dornowski and Perzyna 2002, 2006). Furthermore, on continuum level, the initial porosity exists in metals (denoted by ξ_0) which is in the order of $\xi_0 \approx 10^{-4} \div 10^{-3}$ (Nemes and Eftis 1991).

Definition of Material Functions for Adiabatic Process

Evolution Equations for Internal State Variables

The evolution equations for the internal state variables are postulated in the following rate form:

$$\mathbf{d}^p = \Lambda \mathbf{p}, \quad (63)$$

$$\mathbf{L}_v \boldsymbol{\xi} = \Lambda^h \frac{\partial h^*}{\partial \boldsymbol{\tau}} + \Lambda^g \frac{\partial g^*}{\partial \boldsymbol{\tau}}, \quad (64)$$

where Λ , Λ^h , and Λ^g define the intensity of the viscoplastic flow, the microdamage nucleation, and the microdamage growth, respectively, while \mathbf{p} , $\frac{\partial h^*}{\partial \boldsymbol{\tau}}$, and $\frac{\partial g^*}{\partial \boldsymbol{\tau}}$ define viscoplastic flow direction, microdamage nucleation direction, and microdamage growth direction, respectively.

Material Functions

Thermoelastic Range

Because of existence of the initial microdamage state, the microdamage tensor also influences thermoelastic range. Mathematically it is expressed by putting the stiffness operator to be a function of the microdamage state (Sumelka and Glema 2007, 2008):

$$\mathcal{L}^e = \mathcal{L}^e(\boldsymbol{\xi}). \quad (65)$$

The explicit definition of Eq. (65) is given under the assumption that the operator \mathcal{L}^e in damage directions is coaxial with principal directions of microdamage tensor $\boldsymbol{\xi}$, namely,

$$\begin{aligned} \hat{\mathcal{L}}_{ijklm}^e = & (\hat{\mathcal{C}})_{ijklm} \delta_{ij} \delta_{km} + (\hat{\mathcal{D}})_{ijklm} (\delta_{ik} \delta_{jm} + \delta_{im} \delta_{jk}) + \\ & (\hat{\mathcal{E}})_{ijklm} (\delta_{ik} \delta_{jm} - \delta_{im} \delta_{jk}), \end{aligned} \quad (66)$$

where \mathcal{C} , \mathcal{D} , and \mathcal{E} are the fourth-order material parameters tensors; $(\hat{\mathcal{C}})_{ijklm}$, $(\hat{\mathcal{D}})_{ijklm}$, and $(\hat{\mathcal{E}})_{ijklm}$ are projections of those tensors onto the orthonormal basis vectors \hat{e}_i (Holzapfel 2000) (alternatively we could use the notation introduced in Mura 1987, e.g., instead of $(\hat{\mathcal{C}})_{ijklm}$, we could write $\hat{\mathcal{C}}_{IJKM}$ with the annotation that uppercase indices take on the same numbers as the corresponding lowercase ones but are not summed up – see also Skolnik et al. 2008.); δ denotes the Kronecker symbol; and $(\hat{\cdot})$ denotes damage directions. The tensors in Eq. (66) which describe the material parameters are defined in the following way:

$$\hat{\mathcal{C}}_{ijklm} = (\hat{\mathcal{G}})_{ijklm}, \quad (67)$$

$$\hat{\mathcal{D}}_{ijklm} = (\hat{\mathcal{G}})_{ijklm} \frac{1}{2(1 - 2(\hat{\mathbf{v}})_{ijklm})}, \quad (68)$$

$$\hat{\mathcal{E}}_{ijklm} = (\hat{\mathcal{G}})_{ijklm} \frac{1 - 4(\hat{\mathbf{v}})_{ijklm}}{2(1 - 2(\hat{\mathbf{v}})_{ijklm})}, \quad (69)$$

where

$$\hat{\mathcal{G}}_{ijklm} = G_0(1 - (\hat{\mathbf{Y}})_{ijklm}) \left(1 - \frac{6K_0 + 12G_0}{9K_0 + 8G_0} (\hat{\mathbf{Y}})_{ijklm} \right), \quad (70)$$

$$\hat{\mathcal{K}}_{ijklm} = \frac{4G_0K_0(1 - (\hat{\mathbf{Y}})_{ijklm})}{4G_0 + 3K_0(\hat{\mathbf{Y}})_{ijklm}}, \quad (71)$$

$$\hat{\mathbf{v}}_{ijklm} = \frac{1}{2} \frac{3\hat{\mathcal{K}}_{ijklm} - 2\hat{\mathcal{G}}_{ijklm}}{3\hat{\mathcal{K}}_{ijklm} + \hat{\mathcal{G}}_{ijklm}}. \quad (72)$$

The parameters G_0 and K_0 are the shear and bulk modulus of the material matrix, respectively, and $\hat{\mathbf{Y}}$ is defined as

$$\hat{\mathbf{Y}}_{ijklm} = \sqrt{3} \sqrt{\hat{\xi}_{ij} \hat{\xi}_{km}} \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j \otimes \hat{\mathbf{e}}_k \otimes \hat{\mathbf{e}}_m. \quad (73)$$

Finally, applying Voigt notation, the matrix representation of the constitutive tensor $\hat{\mathcal{L}}^e$ is

$$\hat{\mathcal{L}}^e = \begin{bmatrix} \hat{\mathcal{L}}^e(\hat{\mathbf{Y}}_{1111}) & \hat{\mathcal{L}}^e(\hat{\mathbf{Y}}_{1122}) & \hat{\mathcal{L}}^e(\hat{\mathbf{Y}}_{1133}) & 0 & 0 & 0 \\ \hat{\mathcal{L}}^e(\hat{\mathbf{Y}}_{2211}) & \hat{\mathcal{L}}^e(\hat{\mathbf{Y}}_{2222}) & \hat{\mathcal{L}}^e(\hat{\mathbf{Y}}_{2233}) & 0 & 0 & 0 \\ \hat{\mathcal{L}}^e(\hat{\mathbf{Y}}_{3311}) & \hat{\mathcal{L}}^e(\hat{\mathbf{Y}}_{3322}) & \hat{\mathcal{L}}^e(\hat{\mathbf{Y}}_{3333}) & 0 & 0 & 0 \\ 0 & 0 & 0 & \hat{\mathcal{L}}^e(\hat{\mathbf{Y}}_{1212}) & 0 & 0 \\ 0 & 0 & 0 & 0 & \hat{\mathcal{L}}^e(\hat{\mathbf{Y}}_{2323}) & 0 \\ 0 & 0 & 0 & 0 & 0 & \hat{\mathcal{L}}^e(\hat{\mathbf{Y}}_{3131}) \end{bmatrix}. \quad (74)$$

Notice that, $\hat{\mathbf{Y}}_{1212} = \hat{\mathbf{Y}}_{2323} = \hat{\mathbf{Y}}_{3131} = 0$. It is important that for $\hat{\xi} \rightarrow \hat{\mathbf{0}}$ we arrive at

$$\lim_{\hat{\xi} \rightarrow \hat{\mathbf{0}}} \hat{\mathcal{G}}_{ijklm} \equiv G_0,$$

$$\lim_{\hat{\xi} \rightarrow \hat{\mathbf{0}}} \hat{\mathcal{K}}_{ijklm} \equiv K_0, \quad (75)$$

$$\lim_{\hat{\xi} \rightarrow \hat{\mathbf{0}}} \hat{\mathbf{v}}_{ijklm} \equiv \frac{1}{2} \frac{3K_0 - 2G_0}{3K_0 + G_0}.$$

Thus, the tensor \mathcal{L}^e becomes the well-known (isotropic) Hooke's elastic one. For $\hat{\xi} \rightarrow \hat{\mathbf{1}}$, we obtain material annihilation:

$$\begin{aligned} \lim_{\hat{\xi} \rightarrow \hat{\mathbf{1}}} \hat{G}_{ijkl} &\equiv 0, \\ \lim_{\hat{\xi} \rightarrow \hat{\mathbf{1}}} \hat{K}_{ijkl} &\equiv 0, \\ \lim_{\hat{\xi} \rightarrow \hat{\mathbf{1}}} \hat{v}_{ijkl} &\equiv \frac{45K_0^2 + 16G_0^2 + 48K_0G_0}{117K_0^2 - 16G_0^2 + 96K_0G_0}. \end{aligned} \quad (76)$$

For isotropic microdamage, when we substitute the microdamage tensor by scalar damage parameter, the Mackenzie formulation is obtained (Mackenzie 1950; Perzyna 1986a).

Nonetheless, in following discussion, it is assumed that the elastic range is isotropic and independent of the microdamage state; thus, elastic constitutive tensor \mathcal{L}^e takes the form

$$\mathcal{L}^e = 2\mu\mathcal{I} + \lambda(\mathbf{g} \otimes \mathbf{g}), \quad (77)$$

where μ and λ are Lamé constants. Such assumption is justified because of small influence of elastic range in the extremely dynamic processes.

Because of above assumption, the thermal expansion effects are also isotropic; thus, thermal operator \mathcal{L}^{th} is

$$\mathcal{L}^{th} = (2\mu + 3\lambda)\theta\mathbf{g}, \quad (78)$$

where θ is thermal expansion coefficient.

Viscoplastic Range

We accept common assumptions for TTV theory concerning the rate of viscoplastic strains \mathbf{d}^p Perzyna (1963, 1966), namely,

$$\Lambda^{vp} = \frac{1}{T_m} \left\langle \Phi^{vp} \left(\frac{f}{\kappa} - 1 \right) \right\rangle = \frac{1}{T_m} \left\langle \left(\frac{f}{\kappa} - 1 \right)^{m_{pl}} \right\rangle, \quad (79)$$

$$f = \left\{ J_2' + \left[n_1(\vartheta) + n_2(\vartheta)(\boldsymbol{\xi} : \boldsymbol{\xi})^{\frac{1}{2}} \right] J_1^2 \right\}^{\frac{1}{2}}, \quad (80)$$

$$n_1(\vartheta) = 0, \quad n_2(\vartheta) = n = \text{const.}, \quad (81)$$

$$\kappa = \{ \kappa_s(\vartheta) - [\kappa_s(\vartheta) - \kappa_0(\vartheta)] \exp[-\delta(\vartheta) \in^p] \} \left[1 - \left(\frac{(\boldsymbol{\xi} : \boldsymbol{\xi})^{\frac{1}{2}}}{\xi^F} \right)^{\beta(\vartheta)} \right], \quad (82)$$

$$\begin{aligned} \bar{\vartheta} &= \frac{\vartheta - \vartheta_0}{\vartheta_0}, \quad \kappa_s(\vartheta) = \kappa_s^* - \kappa_s^{**} \bar{\vartheta}, \quad \kappa_0(\vartheta) = \kappa_0^* - \kappa_0^{**} \bar{\vartheta}, \\ \delta(\vartheta) &= \delta^* - \delta^{**} \bar{\vartheta}, \quad \beta(\vartheta) = \beta^* - \beta^{**} \bar{\vartheta}. \end{aligned} \quad (83)$$

$$\mathbf{p} = \frac{\partial f}{\partial \boldsymbol{\tau}} \Big|_{\xi = \text{const}} \left(\left\| \frac{\partial f}{\partial \boldsymbol{\tau}} \right\| \right)^{-1} = \frac{1}{[2J'_2 + 3A^2(\text{tr}\boldsymbol{\tau})^2]^{\frac{1}{2}}} [\boldsymbol{\tau}' + A \text{tr}\boldsymbol{\tau} \boldsymbol{\delta}], \quad (84)$$

where yield surface f is considered in the form suggested in Shima and Oyane (1976), Perzyna (1986a,b) and Glema et al. (2009), κ is the work hardening-softening function (Perzyna 1986a; Nemes and Eftis 1993), $\boldsymbol{\tau}'$ represents the stress deviator, J_1 and J'_2 are the first and the second invariants of Kirchhoff stress tensor and deviatoric part of the Kirchhoff stress tensor, respectively, and $A = 2(n_1 + n_2(\boldsymbol{\xi} : \boldsymbol{\xi})^{\frac{1}{2}})$.

Microdamage Nucleation and Growth

The material functions for the microdamage tensor evolution are Dornowski (1999) and Glema et al. (2009) as follows:

- for the anisotropic microdamage nucleation effects,

$$\frac{\partial h^*}{\partial \boldsymbol{\tau}} = \left\langle \frac{\partial \hat{h}}{\partial \boldsymbol{\tau}} \right\rangle \left\| \left\langle \frac{\partial \hat{h}}{\partial \boldsymbol{\tau}} \right\rangle \right\|^{-1}, \quad (85)$$

$$\hat{h} = \frac{1}{2} \mathbf{e} : \mathcal{N} : \mathbf{e}, \quad \mathcal{N} = \mathcal{I}^s, \quad \mathcal{N}_{ijkl} = \frac{1}{2} (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}), \quad (86)$$

$$I_n = a_1 J_1 + a_2 (J'_2)^{\frac{1}{2}} + a_3 (J'_3)^{\frac{1}{3}}, \quad (87)$$

$$\tau_n = \left(1 - (\boldsymbol{\xi} : \boldsymbol{\xi})^{\frac{1}{2}} \right) (\tau_n^* - \tau_n^{**}(\vartheta) + \tau_n^{***}(\dot{\epsilon}^p)), \quad (88)$$

$$\tau_n^{**}(\vartheta) = \tau_n^{**} \frac{\vartheta - \vartheta_0}{\vartheta_0}, \quad \tau_n^{***}(\dot{\epsilon}^p) = \tau_n^{***} \log \frac{\dot{\epsilon}^p - \dot{\epsilon}_Q^p}{\dot{\epsilon}_Q^p}, \quad (89)$$

$$\Phi^n \left(\frac{I_n}{\tau_n} - 1 \right) = \left(\frac{I_n}{\tau_n} - 1 \right)^{m_n}. \quad (90)$$

where \bar{a}_i ($i = 1, 2, 3$) are the material parameters, J'_3 is the third invariant of deviatoric part of the Kirchhoff stress tensor, and τ_n is the void nucleation threshold stress.

- for the anisotropic microdamage growth effects,

$$\frac{\partial g^*}{\partial \boldsymbol{\tau}} = \left\langle \frac{\partial \hat{g}}{\partial \boldsymbol{\tau}} \right\rangle \left\| \left\langle \frac{\partial \hat{g}}{\partial \boldsymbol{\tau}} \right\rangle \right\|^{-1}, \quad (91)$$

$$\hat{g} = \frac{1}{2} \boldsymbol{\tau} : \mathcal{G} : \boldsymbol{\tau}, \quad \mathcal{G} = \mathcal{I}^s, \quad \mathcal{G}_{ijkl} = \frac{1}{2} (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}), \quad (92)$$

$$I_g = b_1 J_1 + b_2 (J'_2)^{\frac{1}{2}} + b_3 (J'_3)^{\frac{1}{3}}, \quad (93)$$

$$\tau_{eq} = \tau_{eq}(\boldsymbol{\xi}, \vartheta, \epsilon^p) = c(\vartheta) \left(1 - (\boldsymbol{\xi} : \boldsymbol{\xi})^{\frac{1}{2}}\right) \ln \frac{1}{(\boldsymbol{\xi} : \boldsymbol{\xi})^{\frac{1}{2}}} \\ \{2\kappa_s(\vartheta) - [\kappa_s(\vartheta) - \kappa_0(\vartheta)] F(\xi_0, \boldsymbol{\xi}, \vartheta)\}, \quad (94)$$

$$F = F(\xi_0, \boldsymbol{\xi}, \vartheta) = \left(\frac{\xi_0}{1 - \xi_0} \frac{1 - (\boldsymbol{\xi} : \boldsymbol{\xi})^{\frac{1}{2}}}{(\boldsymbol{\xi} : \boldsymbol{\xi})^{\frac{1}{2}}}\right)^{\frac{2}{3}\delta(\vartheta)} + \left(\frac{1 - (\boldsymbol{\xi} : \boldsymbol{\xi})^{\frac{1}{2}}}{1 - \xi_0}\right)^{\frac{2}{3}\delta(\vartheta)}, \quad (95)$$

$$\Phi^g \left(\frac{I_g}{\tau_{eq}} - 1\right) = \left(\frac{I_g}{\tau_{eq}} - 1\right)^{m_g}, \quad (96)$$

where τ_{eq} is the void growth threshold stress and \bar{b}_i ($i = 1, 2, 3$) are the material parameters.

- for thermal effects,

$$\operatorname{div} \mathbf{q} = \mathbf{0}, \quad (97)$$

$$\mathbf{k} = \boldsymbol{\tau}. \quad (98)$$

It should be noticed that for the microdamage mechanism, it is assumed that Dornowski (1999) and Glema et al. (2009):

- velocity of the microdamage nucleation is coaxial with the principal directions of strain state,
- velocity of the microdamage growth is coaxial with the principal directions of stress state,
- only positive (tension) principal strain (stress) induces the nucleation (growth) of the microdamage.

Concluding, for an adiabatic process, one needs to identify 36 material parameters:

- 5 parameters related to viscoplastic evolution,
- 17 parameters related to anisotropic microdamage evolution (nucleation 8, growth 5, fracture porosity 4),
- 7 parameters shared with viscoplastic and microdamage evolution,
- 4 parameters related to thermal evolution,
- 2 elasticity parameters, and
- reference density.

Numerical Examples: Modeling of Spalling Phenomena

Introductory Remarks

Herein the results discussed in Łodygowski and Sumelka (2015) are developed to map the damage nucleation effects, also. From the TTV properties like: (i) the description is invariant with respect to any diffeomorphism (covariant material model), (ii) the obtained evolution problem is well posed, (iii) sensitivity to the rate of deformation, (iv) finite elasto-viscoplastic deformations, (v) plastic non-normality, (vi) dissipation effects, (vii) thermomechanical couplings, and (viii) length scale sensitivity; the last one plays herein the exposed role. Namely, precise modeling of the geometry and intensity of the localized deformation zones, the initiation time of macrodamage, its direction, and the final fracture pattern are aimed to be modeled in details.

A series of 3D tension analyses were included to understand the influence of mesh density, element type, and/or the relaxation time on numerical results – cf. Figs. 4, 5, 6, and 7. Fully integrated elements are the most proper choice considering damage analysis where damaged elements are removed from mesh causing a considerable change in the numerical model. The reduced elements in which single integration point exists are more sensitive – herein damage of an element is controlled in a single point whereas in a full integrated element in eight points, damage criterion must be fulfilled to delete the element. Furthermore, the relaxation time controls truly the rate of viscoplastic deformation and at the same time the initiation and growth of damage. Nonetheless, because of extreme dynamic regime (the velocity of tension was $60 \frac{\text{m}}{\text{s}}$ and the specimen length c.a. 20 mm), the shear band zones were localized close to loaded end.

In the remaining part of this section before analysis of the spalling phenomenon, initial boundary value problem (IBVP) is emphasized together with some comments on computer implementation of TTV in the Abaqus/Explicit finite element program.

Initial Boundary Value Problem

To solve the thermomechanical problem in terms of the presented TTV means to find ϕ , \mathbf{v} , ρ , τ , ξ , and ϑ as functions of t and position \mathbf{x} such that the following equations are satisfied Perzyna (1994), Łodygowski (1996) and Łodygowski and Perzyna (1997a,b):

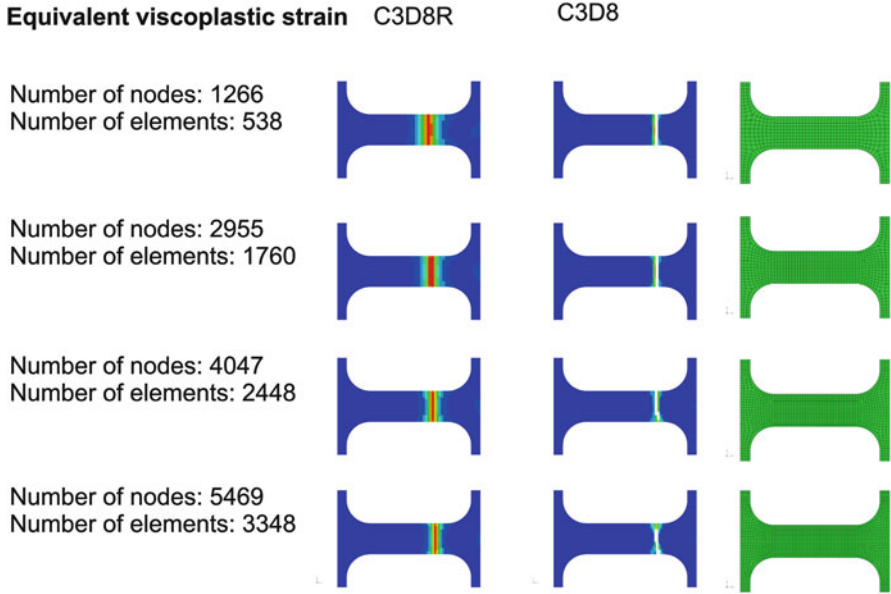


Fig. 4 The comparison of equivalent plastic strain localization zones vs. element type and mesh size

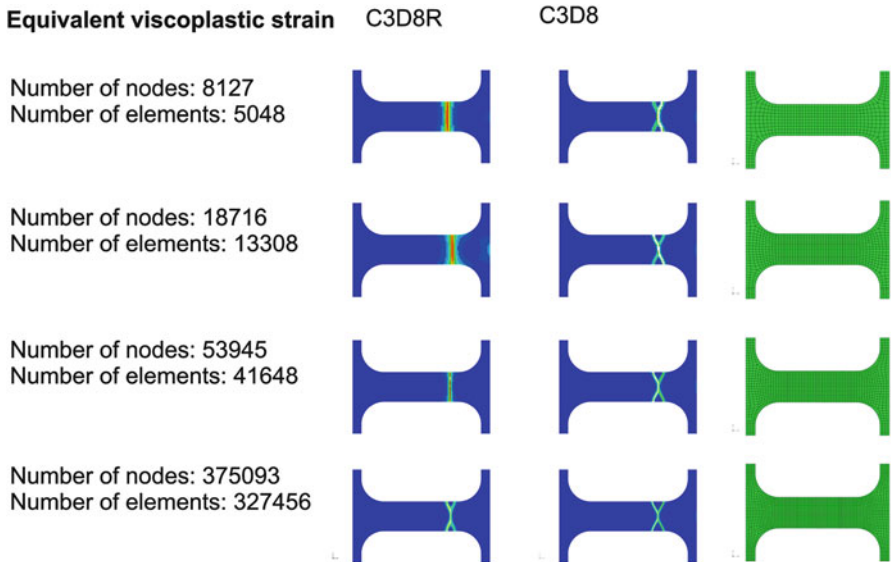


Fig. 5 The comparison of equivalent plastic strain localization zones vs. element type and mesh size

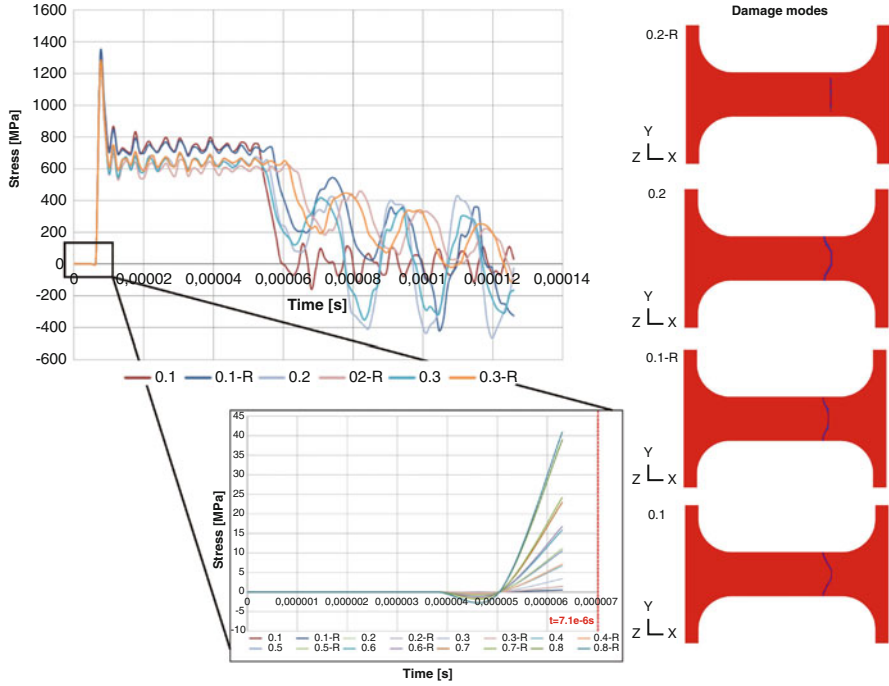


Fig. 6 The comparison of strain and time at the fixed end curves dependently on element type and mesh size

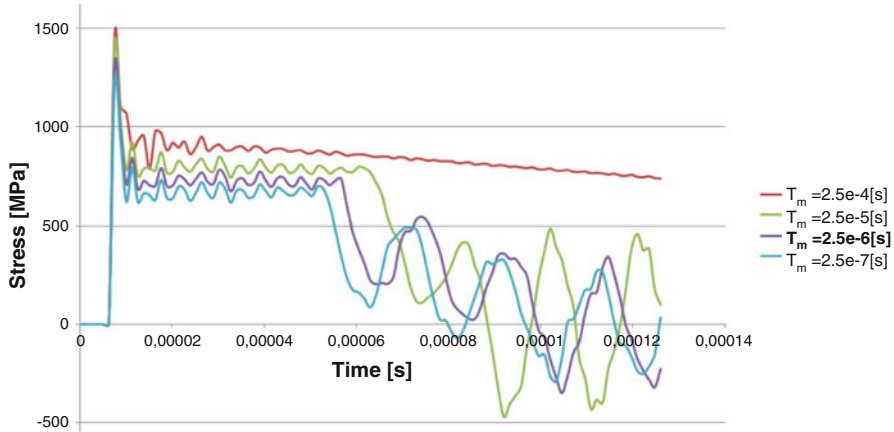


Fig. 7 The comparison of strain and time at the fixed end curves as a function of relaxation time

(i) the field equations:

$$\begin{aligned} \dot{\phi} &= \mathbf{v}, \\ \dot{\mathbf{v}} &= \frac{1}{\rho_{\text{Ref}}} \left(\text{div} \boldsymbol{\tau} + \frac{\boldsymbol{\tau}}{\rho} \cdot \text{grad} \rho - \frac{\boldsymbol{\tau}}{1 - (\boldsymbol{\xi} : \boldsymbol{\xi})^{\frac{1}{2}}} \text{grad} (\boldsymbol{\xi} : \boldsymbol{\xi})^{\frac{1}{2}} \right), \\ \dot{\rho} &= -\rho \text{div} \mathbf{v} + \frac{\rho}{1 - (\boldsymbol{\xi} : \boldsymbol{\xi})^{\frac{1}{2}}} (\mathbf{L}_v \boldsymbol{\xi} : \mathbf{L}_v \boldsymbol{\xi})^{\frac{1}{2}}, \\ \dot{\boldsymbol{\tau}} &= \mathcal{L}^e : \mathbf{d} + 2\boldsymbol{\tau} \cdot \mathbf{d} - \mathcal{L}^{th} \dot{\vartheta} - (\mathcal{L}^e + \mathbf{g}\boldsymbol{\tau} + \boldsymbol{\tau}\mathbf{g}) : \mathbf{d}^p, \\ \dot{\boldsymbol{\xi}} &= 2\boldsymbol{\xi} \cdot \mathbf{d} + \frac{\partial g^*}{\partial \boldsymbol{\tau}} \frac{1}{T_m} \left\langle \Phi^g \left[\frac{I_g}{\tau_{eq}(\boldsymbol{\xi}, \vartheta, \epsilon^p)} - 1 \right] \right\rangle \\ &\quad + \frac{\partial h^*}{\partial \boldsymbol{\tau}} \frac{1}{T_m} \left\langle \Phi^n \left[\frac{I_n}{\tau_n(\boldsymbol{\xi}, \vartheta, \epsilon^p)} - 1 \right] \right\rangle, \\ \dot{\vartheta} &= \frac{\chi^*}{\rho c_p} \boldsymbol{\tau} : \mathbf{d}^p + \frac{\chi^{**}}{\rho c_p} \mathbf{k} : \mathbf{L}_v \boldsymbol{\xi}, \end{aligned} \tag{99}$$

(ii) the boundary conditions:

- (a) displacement ϕ is prescribed on a part Γ_ϕ of $\Gamma(\mathcal{B})$ and tractions $(\boldsymbol{\tau} \cdot \mathbf{n})^a$ are prescribed on a part Γ_τ of $\Gamma(\mathcal{B})$, where $\Gamma_\phi \cap \Gamma_\tau = \emptyset$ and $\Gamma_\phi \cup \Gamma_\tau = \Gamma(\mathcal{B})$,
- (b) heat flux $\mathbf{q} \cdot \mathbf{n} = 0$ is prescribed on $\Gamma(\mathcal{B})$,

(iii) and the initial conditions $\phi, \mathbf{v}, \rho, \boldsymbol{\tau}, \boldsymbol{\xi}$, and ϑ are given for each particle $\mathbf{X} \in \mathcal{B}$ at $t = 0$,

are satisfied.

In Eq. (99)₆, because of adiabatic regime assumption, the first two terms in temperature evolution law Eq. (51) are omitted. It should be emphasized that adiabatic condition assumption weakens the robustness of modeling due to the fact that the first term in Eq. (51) introduces in a natural way the nonlocality. Nevertheless, recall that in the viscoplasticity, the nonlocality results implicitly from the relaxation time parameter (T_m).

This evolution form is even better seen when the operator equation is involved (Łodygowski 1995). Hence, the inhomogeneous abstract Cauchy problem can be formulated in the following form:

$$\dot{\boldsymbol{\varphi}}(t, \mathbf{x}) = \mathcal{A}(t, \mathbf{x}) \cdot \boldsymbol{\varphi}(t, \mathbf{x}) + \mathbf{f}(t, \mathbf{x}, \boldsymbol{\varphi}) \quad \text{for } t \in (0, T] \quad \text{and } \mathbf{x} \in \Omega \tag{100}$$

with the initial condition

$$\boldsymbol{\varphi}(0, \mathbf{x}) = \boldsymbol{\varphi}_0. \tag{101}$$

where

$$\boldsymbol{\varphi} = \begin{bmatrix} \phi \\ \mathbf{v} \\ \rho \\ \boldsymbol{\tau} \\ \boldsymbol{\xi} \\ \vartheta \end{bmatrix}, \quad (102)$$

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\boldsymbol{\tau} \text{grad}}{\rho_{\text{Ref}} \rho} & \frac{\text{div}}{\rho_{\text{Ref}}} & \frac{-\boldsymbol{\tau} \text{grad}}{\rho_{\text{Ref}}(1-\xi)} & 0 \\ 0 & -\rho \text{div} & 0 & 0 & 0 & 0 \\ 0 & \mathcal{L}^e : \text{sym} \frac{\partial}{\partial \mathbf{x}} + 2 \text{sym}(\boldsymbol{\tau} \cdot \frac{\partial}{\partial \mathbf{x}}) & 0 & 0 & 0 & 0 \\ 0 & 2 \text{sym}(\boldsymbol{\xi} \cdot \frac{\partial}{\partial \mathbf{x}}) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad (103)$$

$$\mathbf{f} = \begin{bmatrix} \mathbf{v} \\ \mathbf{0} \\ \frac{\rho}{1-\xi} (\boldsymbol{\Xi} : \boldsymbol{\Xi})^{\frac{1}{2}} \\ - \left[\left(\frac{\chi^*}{\rho c_p} \mathcal{L}^{th} \boldsymbol{\tau} + \mathcal{L}^e + \mathbf{g} \boldsymbol{\tau} + \boldsymbol{\tau} \mathbf{g} \right) : \mathbf{p} \right] \frac{1}{T_m} \left(\left(\frac{f}{\kappa} - 1 \right)^{m_{pl}} \right) - \frac{\chi^{**}}{\rho c_p} \mathbf{k} : \boldsymbol{\Xi} \\ \boldsymbol{\Xi} \\ \frac{\chi^*}{\rho c_p} \boldsymbol{\tau} : \mathbf{p} \frac{1}{T_{m_{pl}}} \left(\left(\frac{f}{\kappa} - 1 \right)^m \right) + \frac{\chi^{**}}{\rho c_p} \mathbf{k} : \boldsymbol{\Xi} \end{bmatrix}, \quad (104)$$

where $\boldsymbol{\Xi} = \mathbf{L}_v \boldsymbol{\xi}$.

Now, it is necessary to split the discussion into two branches. One is so called the abstract Cauchy problem (ACP), which takes into consideration only the initial conditions, and the other one more significant for responsible computations initial boundary value problem (IBVP). For detailed description and discussion, please refer to the works Łodygowski et al. (1994), Łodygowski (1995, 1996), and Łodygowski and Perzyna (1997a). The well-posedness of the ACP or IBVP means that the solution for the components of vector $\boldsymbol{\varphi}$ exists, is unique, and is stable in Hadamard or Lyapunov sense. It seems to be possible to prove the conditions for well-posedness of ACP and, using the Lax theorem, define the conditions for finite element approximation. Unfortunately, for IBVP until now it is extremely difficult in a strict mathematical way to prove the well-posedness of the system of governing equations (for 1D case it was done by Ionescu and Sofonea 1993). One cannot count for the general proof for arbitrary IBVP. Fortunately, we are not completely helpless. Using the numerical computations, one is able to discover the pathological mesh dependency (using different meshes, e.g., more dense, the typical convergence is not achieved) which is for softening behavior the implicit proof that the governing set of equations is not well posed. For this purpose, any finite element computational results have to be verified for different meshes including specific refinement in the localized areas.

Table 1 Material parameters for HSLA-65 steel

$\lambda = 121.154 \text{ GPa}$	$\mu = 80.769 \text{ GPa}$	$\chi^* = 0.8$	$\chi^{**} = 0.1$
$\kappa_s^* = 570 \text{ MPa}$	$\kappa_s^{**} = 129 \text{ MPa}$	$\kappa_0^* = 457 \text{ MPa}$	$\kappa_0^{**} = 103 \text{ MPa}$
$\beta^* = 11.0$	$\beta^{**} = 2.5$	$\delta^* = 6.0$	$\delta^{**} = 1.4$
$a_1 = 0.7$	$a_2 = 0.85$	$a_3 = 0$	$c = 0.067$
$b_1 = 0.02$	$b_2 = 0.5$	$b_3 = 0$	$m_g = 1$
$\ L_v \xi_c\ - s^{-1}$	$\xi^{F*} = 0.36$	$\xi^{F**} = -$	$m_n = 1$
$\tau_n^* = 671 \text{ MPa}$	$\tau_n^{**} = 265 \text{ MPa}$	$\tau_n^{***} = 38 \text{ MPa}$	$m_F -$
$c_p = 470 \text{ J/kgK}$	$\theta = 10^{-6} \text{ K}^{-1}$	$\epsilon_O^p = 10^{-5} \text{ s}^{-1}$	$m_{pl} = 0.14$
$\rho_{\text{Ref}} = 7800 \text{ kg/m}^3$	$n_1 = 0$	$n_2 = 0.25$	$T_m = 2.5 \mu\text{s}$

Material Parameter Identification for HSLA-65 Steel

As in Łodygowski and Sumelka (2015), the HSLA-65 steel experimentally analyzed in Nemat-Nasser and Guo (2005) is considered. Herein, precise identification of such a great number of material parameters is even more troublesome, because of damage nucleation term. Thus parameters presented in Table 1 should be thought as a compromise; hence, small fluctuations of them are possible (dependently on detailed experimental results showing competition of fundamental variables, e.g., temperature, viscoplastic strain, microdamage). For a proposition of a method to reduce the number of material parameters using soft computing methods cf. (Sumelka and Łodygowski 2013).

Computer Implementation in Abaqus/Explicit

The Abaqus/Explicit commercial finite element code has been adapted as a solver. The Abaqus/ Explicit utilizes central difference time integration rule along with the diagonal (“lumped”) element mass matrices. To remove damaged elements from the mesh (elements in which for every integration point fracture porosity was reached, or equivalently load carrying capacity is zero $\kappa \rightarrow 0$) the so-called element deletion method is applied (Song et al. 2008). The model has been implemented in the software, by taking advantage of a user subroutine VUMAT, which is coupled with the Abaqus system (Abaqus 2012).

Because of Lie derivative application in TTV, special attention is paid to the stress update in VUMAT. Namely, the Green-Naghdi rate is calculated by default in Abaqus/Explicit VUMAT user subroutine; thus, the following formula was applied to enforce spatial covariance (Sumelka 2009):

$$\tilde{\tau} \mid_{i+1} = \mathbf{R}^T \mid_{i+1} [\tau \mid_i + \Delta t (2\tau \mid_i \cdot \mathbf{d} \mid_i + L_v \tau \mid_i) + \Upsilon \mid_i] \mathbf{R} \mid_{i+1}, \quad (105)$$

where $\Upsilon \mid_i = -\Delta t (\mathbf{\Omega} \mid_i \cdot \tau \mid_i - \tau \mid_i \cdot \mathbf{\Omega} \mid_i)$ and $\tau \mid_i = \mathbf{R} \mid_i \tilde{\tau} \mid_i \mathbf{R}^T \mid_i$, where $\mathbf{\Omega}$ is the spin tensor. The importance of the subject of matter lays in fact that

the applied objective rate can influence the results considerably – (cf. Sumelka 2013). Furthermore, some of them (e.g., Zaremba 1903) can lead to the nonphysical solutions (Dienes 1979; Lehmann 1972; Nagtegaal and de Jong 1982; Xiao et al. 1997).

Spall Fracture Phenomenon Modeling

In this section, the results presented in Łodygowski and Sumelka (2015) (as mentioned) are resolved including the additional effects which come from microdamage nucleation effects – previously in all elements initial porosity was assumed ($\xi_0 = 6 \cdot 10^{-4}$), namely,

$$\xi_0 = \begin{bmatrix} 34.64 \cdot 10^{-5} & 0 & 0 \\ 0 & 34.64 \cdot 10^{-5} & 0 \\ 0 & 0 & 34.64 \cdot 10^{-5} \end{bmatrix}.$$

As before, spalling will be the result of flat plate impact test (Meyers and Aimone 1983; Curran 1987; Klepaczko 1990; Hanim and Klepaczko 1999; Boidin et al. 2006) (cf. Fig. 8). In this test two, plates with high velocities impact each other. As a result of the impact (caused by the flyer plate), a complete or partial separation of the material can appear in the target plate Fig. 8. This is due to tension in the target plate, induced by the interaction of two waves, incident and reflected. In spall zone, the intensive evolution of microdamage appears and changes dependently on velocity of flyer plate. Furthermore, spall test (among others, e.g., Močko and Kowalewski 2011, 2013) is an important technique to analyze strain-stress behavior of a material.

The numerical model mapping the flat plate impact experiment is presented in Fig. 9. The dimensions of flyer plate are diameter $\phi_{\text{fla}} = 114$ mm and thickness $t_{\text{fla}} = 5$ mm, while for target plate diameter is $\phi_{\text{tar}} = 114$ mm and thickness is $t_{\text{tar}} = 10$ mm. Due to the microdamage anisotropy, full spatial modeling was necessary. The numerical analyses were conducted in Abaqus/Explicit code including user subroutine VUMAT, as mentioned. The C3D8R finite elements (eight-node linear brick, reduced integration element) were applied – in total c.a. $3M$ finite elements were used. The initial velocity of the flyer plate was $v_0 = 500 \frac{\text{m}}{\text{s}}$, and the initial temperature 296 K was assumed.

The initial microdamage state needs detailed discussion. Because of lack of the experimental data concerning the initial microdamage distribution in the specimen, it was assumed as follows (cf. Figs. 10 and 11):

- for a random set of elements, $\xi_0 = 0$ was assumed,
- for majority of elements, $\xi_0 = 6.0e - 4$ was prescribed,
- and finally for few elements (randomly chosen), the porosity close to the fractured one was taken into account.

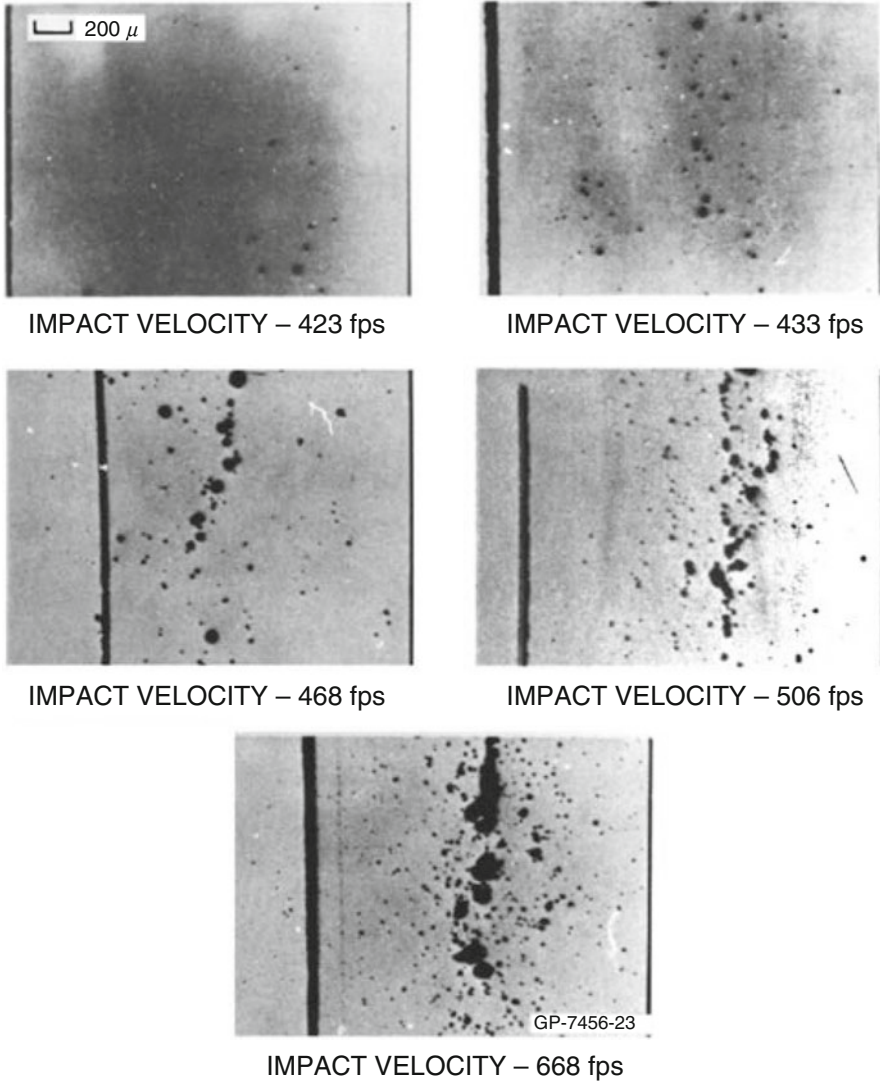


Fig. 8 The final damage of the aluminum 1145 target plate for a constant shot geometry but for different impact velocities (After Barbee et al. 1972)

It is clear that because of tensorial nature of the microdamage state variable, the components of initial microdamage tensor field were provided in every integration point. For simplicity, it was assumed that in the initial configuration the microdamage tensor is isotropic.

Material parameters for TTV model were taken as discussed in section “[Material Parameter Identification for HSLA-65 Steel.](#)”

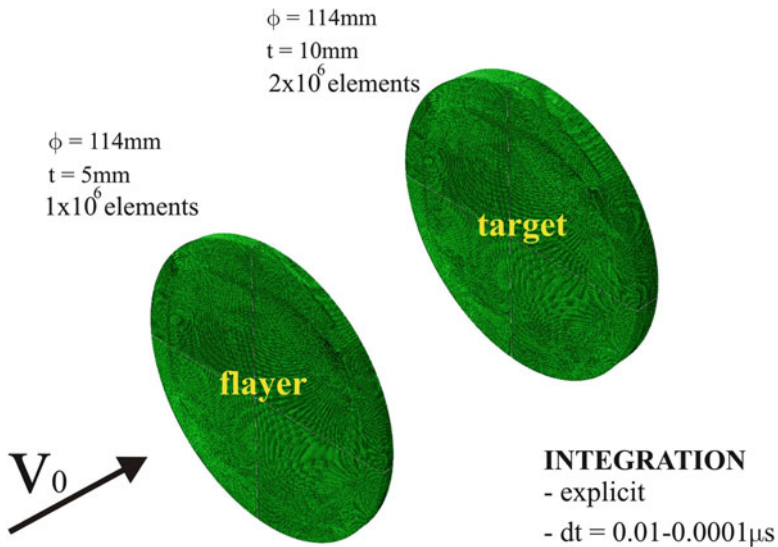


Fig. 9 Numerical model for modelling spall phenomenon

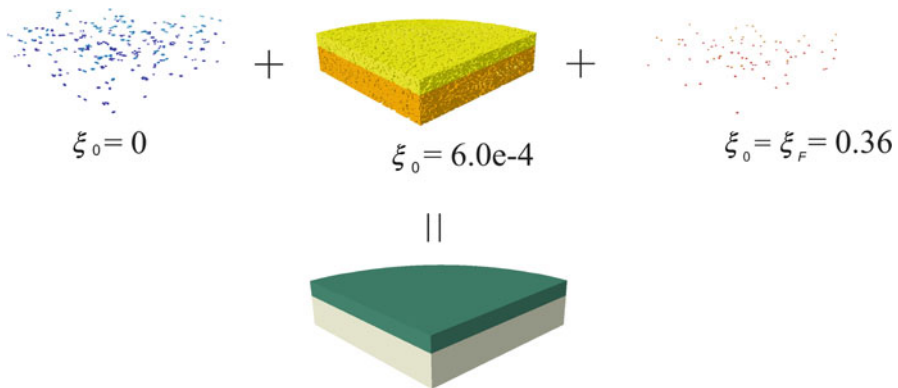


Fig. 10 Initial configuration for spalling analysis: (left) finite elements without initial porosity; (middle) finite elements with initial porosity; (middle) finite elements with initial porosity close to fracture porosity

In Figs. 12 and 13, the numerical results are presented. The initial velocity of flyer plate, equals $v_0 = 500 \frac{\text{m}}{\text{s}}$, was high enough to cause spalling in the target plate (cf. Łodygowski and Sumelka 2015). The sample fields of strain rates, microdamage growth velocity, fracture porosity, and porosity during the process (the analysis of other variables like stresses, thermal stresses, strains, temperatures, etc. is not included) show almost immediate localization of a separation zone as in the experiment. In this sense, the TTV in a presented form is verified and validated.

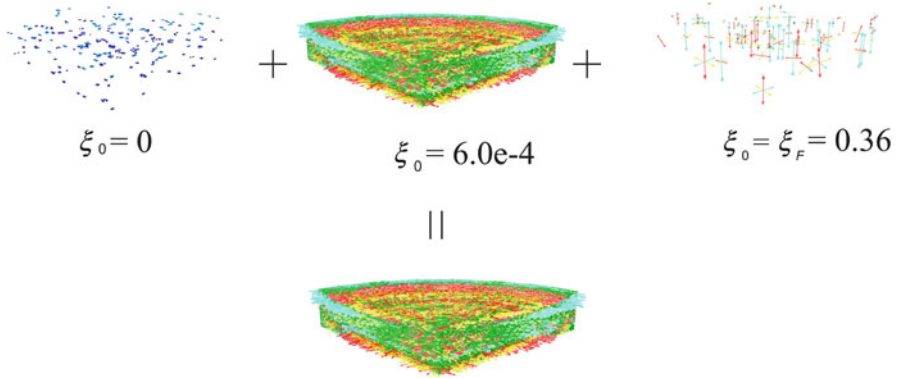


Fig. 11 Initial configuration of a microdamage directions for spalling analysis: (left) finite elements without initial porosity; (middle) finite elements with initially isotropic porosity; (right) finite elements with initial porosity close to fracture porosity

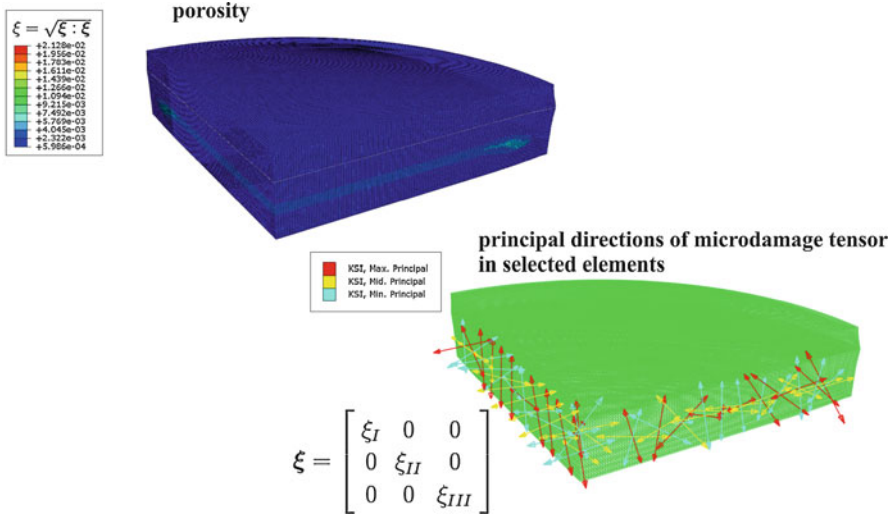


Fig. 12 Comparison of porosity (norm of the microdamage tensor) and principal directions of the microdamage tensor

Conclusions

The presented Perzyna Theory of Thermo-Viscoplasticity (TTV) is motivated by the micromechanical experimental results and ensures unique solution of the posed IBVP. On the other side, the attractiveness of TTV model is reduced due to the high number of material parameters needed for practical applications. Nonetheless, such adverse factor will always appear in the phenomenological modeling if qualitative and quantitative results are needed.

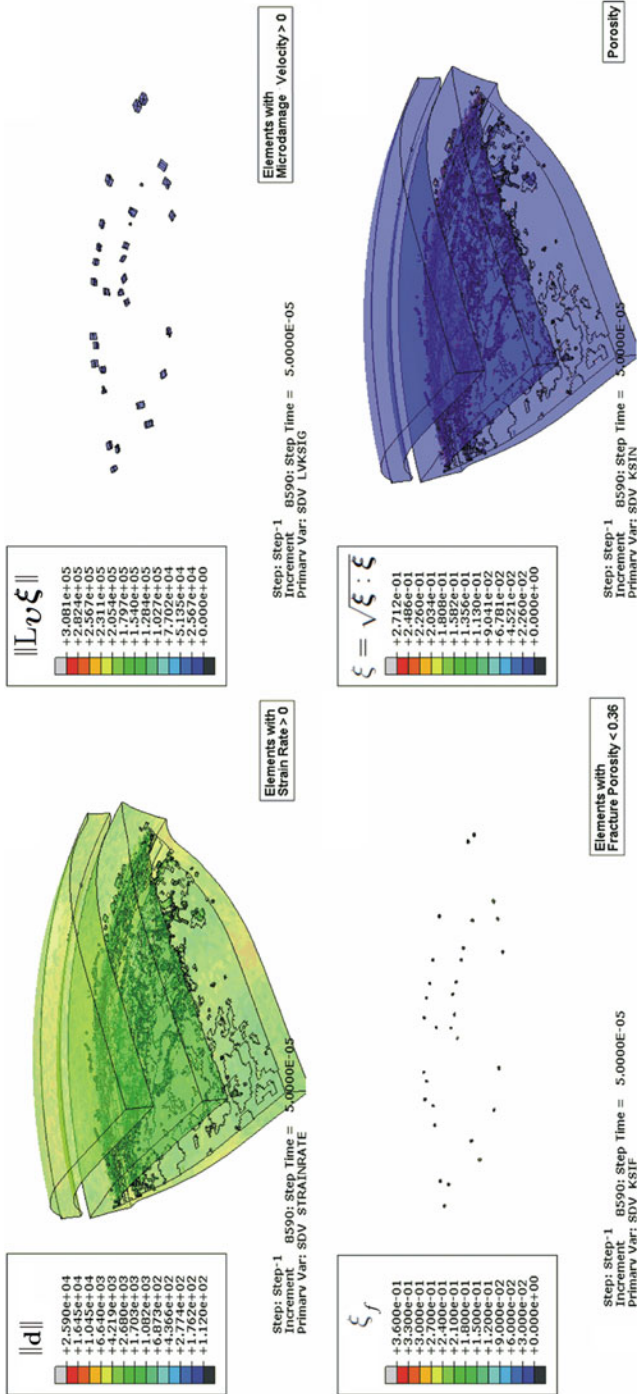


Fig. 13 The plots of strain rates, microdamage growth velocity, fracture porosity, and porosity for time process $5 \cdot 10^{-6}$ s (flayer velocity $v_0 = 550 \frac{m}{s}$)

It is commonly recognized that nonlocal models are more robust than local ones. TTV belongs to the class of implicit nonlocal models, in the sense that there is no direct measure of nonlocal action in the model but some phenomenological material parameters can be interpreted as one that control the scale effects. In TTV this parameter is the relaxation time – deduced from the analysis of the single crystal deformation. As presented, the relaxation time (T_m) can be viewed not only as a microstructural parameter (which includes the information about the characteristic size of plastic deformation) but also as a mathematical regularization parameter.

Finally, it is important to emphasize that today over 50 years after publishing the basic paper on the theory of thermo-viscoplasticity by Perzyna, its concepts are still vivid and stimulate new findings, e.g., *fractional viscoplasticity*.

References

- R. Abraham, J.E. Marsden, T. Ratiu, *Manifolds, Tensor Analysis and Applications* (Springer, Berlin, 1988)
- Abaqus, *Abaqus Version 6.12 Collection* (SIMULIA Worldwide Headquarters, Providence, 2012)
- E.C. Aifantis, Strain gradient interpretation of size effects. *Int. J. Fract.* **95**, 299–314 (1999)
- R.J. Asaro, Crystal plasticity. *J. Appl. Mech.* **50**, 921–934 (1983)
- T.W. Barbee, L. Seaman, R. Crewdson, D. Curran, Dynamic fracture criteria for ductile and brittle metals. *J. Mater.* **7**, 393–401 (1972)
- X. Boidin, P. Chevrier, J.R. Klepaczko, H. Sabar, Identification of damage mechanism and validation of a fracture model based on mesoscale approach in spalling of titanium alloy. *Int. J. Solids Struct.* **43**(14–15), 4029–4630 (2006)
- D.R. Curran, L. Seaman, D.A. Shockey, Dynamic failure of solids. *Phys. Rep.* **147**(5–6), 253–388 (1987)
- S. Cochran, D. Banner, Spall studies in uranium. *J. Appl. Phys.* **48**(7), 2729–2737 (1988)
- R. de Borst, J. Pamin, Some novel developments in finite element procedures for gradient-dependent plasticity. *Int. J. Numer. Methods Eng.* **39**, 2477–2505 (1996)
- J.K. Dienes, On the analysis of rotation and stress rate in deforming bodies. *Acta Mech.* **32**, 217–232 (1979)
- P. Dłużewski, *Continuum Theory of Dislocations as a Theory of Constitutive Modelling of Finite Elastic-Plastic Deformations*. Volume 13 of IFTR Reports. Institute of Fundamental Technological Research – Polish Academy of Science, 1996. (D.Sc. Thesis – in Polish)
- W. Dornowski, Influence of finite deformations on the growth mechanism of microvoids contained in structural metals. *Arch. Mech.* **51**(1), 71–86 (1999)
- W. Dornowski, P. Perzyna, Analysis of the influence of various effects on cycle fatigue damage in dynamic process. *Arch. Appl. Mech.* **72**, 418–438 (2002)
- W. Dornowski, P. Perzyna, Numerical investigation of localized fracture phenomena in inelastic solids. *Found. Civil Environ. Eng.* **7**, 79–116 (2006)
- M.K. Duszek–Perzyna, P. Perzyna, Analysis of the influence of different effects on criteria for adiabatic shear band localization in inelastic solids, vol. 50. *Material Instabilities: Theory and Applications* (ASME, New York, 1994)
- A.C. Eringen, Linear theory of nonlocal elasticity and dispersion of plane-waves. *Int. J. Eng. Sci.* **10**(5), 233–248 (1972a)
- A.C. Eringen, Nonlocal polar elastic continua. *Int. J. Eng. Sci.* **10**, 1–16 (1972b)
- A.C. Eringen, On differential-equations of nonlocal elasticity and solutions of screw dislocation and surface-waves. *J. Appl. Phys.* **54**(9), 4703–4710 (1983)
- N.A. Fleck, J.W. Hutchinson, Strain gradient plasticity. *Adv. Appl. Mech.* **33**, 295–361 (1997)

- M. Frewer, More clarity on the concept of material frame-indifference in classical continuum mechanics. *Acta Mech.* **202**(1–4), 213–246 (2009)
- A. Glema, Analysis of Wave Nature in Plastic Strain Localization in Solids. Volume 379 of *Rozprawy*, Publishing House of Poznan University of Technology, 2004 (in Polish)
- A. Glema, T. Łodygowski, On importance of imperfections in plastic strain localization problems in materials under impact loading. *Arch. Mech.* **54**(5–6), 411–423 (2002)
- A. Glema, W. Kakol, T. Łodygowski, Numerical modelling of adiabatic shear band formation in a twisting test. *Eng. Trans.* **45**(3–4), 419–431 (1997)
- A. Glema, T. Łodygowski, P. Perzyna, Interaction of deformation waves and localization phenomena in inelastic solids. *Comput. Methods Appl. Mech. Eng.* **183**, 123–140 (2000)
- A. Glema, T. Łodygowski, P. Perzyna, Localization of plastic deformations as a result of wave interaction. *Comput. Assist. Mech. Eng. Sci.* **10**(1), 81–91 (2003)
- A. Glema, T. Łodygowski, W. Sumelka, P. Perzyna, The numerical analysis of the intrinsic anisotropic microdamage evolution in elasto-viscoplastic solids. *Int. J. Damage Mech.* **18**(3), 205–231 (2009)
- A. Glema, T. Łodygowski, W. Sumelka, Piotr perzyna – scientific conductor within theory of thermo-viscoplasticity. *Eng. Trans.* **62**(3), 193–219 (2014)
- A.E. Green, R.S. Rivlin, Multipolar continuum mechanics. *Arch. Ration. Mech. Anal.* **17**(2), 113–147 (1964)
- S. Hanim, J.R. Klepaczko, Numerical study of spalling in an aluminum alloy 7020 – T6. *Int. J. Impact Eng.* **22**, 649–673 (1999)
- O.M. Heeres, A.S.J. Suiker, R. de Borst, A comparison between the perzyna viscoplastic model and the consistency viscoplastic model. *Eur. J. Mech. A. Solids* **21**(1), 1–12 (2002)
- R. Hill, Aspects of invariance in solid mechanics. *Adv. Appl. Mech.* **18**, 1–75 (1978)
- G.A. Holzapfel, *Nonlinear Solid Mechanics – A Continuum Approach for Engineering* (Chichester, England, 2000)
- I.R. Ionescu, M. Sofonea, *Functional and Numerical Methods in Viscoplasticity* (Oxford University Press, Oxford/New York/Tokyo, 1993)
- A.A. Kilbas, H.M. Srivastava, J.J. Trujillo, *Theory and Applications of Fractional Differential Equations* (Elsevier, Amsterdam, 2006)
- J.R. Klepaczko, Dynamic crack initiation, some experimental methods and modelling, in *Crack Dynamics in Metallic Materials*, ed. by J.R. Klepaczko (Springer, Vienna, 1990), pp. 255–453
- E. Kröner, On the physical reality of torque stresses in continuum mechanics. *Int. J. Eng. Sci.* **1**, 261–278 (1963)
- Th. Lehmann, Anisotrope plastische Formänderungen. *Romanian J. Tech. Sci. Appl. Mech.* **17**, 1077–1086 (1972)
- T. Łodygowski, On avoiding of spurious mesh sensitivity in numerical analysis of plastic strain localization. *Comput. Assist. Mech. Eng. Sci.* **2**, 231–248 (1995)
- T. Łodygowski, Theoretical and Numerical Aspects of Plastic Strain Localization. Volume 312 of D.Sc. Thesis, Publishing House of Poznan University of Technology, 1996
- T. Łodygowski, P. Perzyna, Localized fracture of inelastic polycrystalline solids under dynamic loading process. *Int. J. Damage Mech.* **6**, 364–407 (1997a)
- T. Łodygowski, P. Perzyna, Numerical modelling of localized fracture of inelastic solids in dynamic loading process. *Int. J. Numer. Methods Eng.* **40**, 4137–4158 (1997b)
- T. Łodygowski, W. Sumelka, Anisotropic damage for extreme dynamics, in *Handbook of Damage Mechanics Nano to Macro Scale for Materials and Structures*, ed. by G.Z. Voyiadjis (Springer, New York, 2015), pp. 1185–1220
- T. Łodygowski, P. Perzyna, M. Lengnick, E. Stein, Viscoplastic numerical analysis of dynamic plastic shear localization for a ductile material. *Arch. Mech.* **46**(4), 541–557 (1994)
- J.K. Mackenzie, The elastic constants of a solids containing spherical holes. *Proc. Phys. Soc.* **63B**, 2–11 (1950)
- J.E. Marsden, T.J.H. Hughes, *Mathematical Foundations of Elasticity* (Prentice-Hall, New Jersey, 1983). <https://www.sciencedirect.com/science/article/pii/0079642583900038>

- M.A. Meyers, C.T. Aimone, Dynamic fracture (Spalling) of materials, *Progress in Material Science*, **28**(1), 1–96 (1983)
- R.D. Mindlin, Micro-structure in linear elasticity. *Arch. Ration. Mech. Anal.* **16**(1), 51–78 (1964)
- R.D. Mindlin, Second gradient of strain and surface-tension in linear elasticity. *Int. J. Solids Struct.* **1**(4), 417–438 (1965)
- R.D. Mindlin, N. Eshel, On first strain-gradient theories in linear elasticity. *Int. J. Solids Struct.* **4**(1), 109–124 (1968)
- R.D. Mindlin, H.F. Tiersten, Effects of couple-stresses in linear elasticity. *Arch. Ration. Mech. Anal.* **11**(5), 415–448 (1962)
- W. Moćko, Z.L. Kowalewski, Mechanical properties of a359/sicp metal matrix composites at wide range of strain rates. *Appl. Mech. Mater.* **82**, 166–171 (2011)
- W. Moćko, Z.L. Kowalewski, Perforation test as an accuracy evaluation tool for a constitutive model of austenitic steel. *Arch. Metall. Mater.* **58**(4), 1105–1110 (2013)
- J.V. Morán, Continuum Models for the Dynamic Behavior of 1D Nonlinear Structured Solids. Doctoral Thesis, Publishing House of the Universidad Carlos III de Madrid, 2016
- T. Mura, *Micromechanics of Defects in Solids* (Kluwer Academic, Dordrecht, 1987)
- J.C. Nagtegaal, J.E. de Jong, Some aspects of non-isotropic work-hardening in finite strain plasticity, in *Proceedings of the Workshop on Plasticity of Metals at Finite Strain: Theory, Experiment and Computation*, ed. by E.H. Lee, R.L. Mallet (Stanford University, 1982), pp. 65–102
- S. Nemat-Nasser, W.-G. Guo, Thermomechanical response of HSLA-65 steel plates: experiments and modeling. *Mech. Mater.* **37**, 379–405 (2005)
- J.A. Nemes, J. Eftis, Several features of a viscoplastic study of plate-impact spallation with multidimensional strain. *Comput. Struct.* **38**(3), 317–328 (1991)
- J.A. Nemes, J. Eftis, Constitutive modelling of the dynamic fracture of smooth tensile bars. *Int. J. Plast.* **9**(2), 243–270 (1993)
- J. Ostrowska-Maciejewska, *Mechanika ciał odkształcalnych* (PWN, Warszawa, 1994)
- S.K. Park, X.-L. Gao, Variational formulation of a modified couple stress theory and its application to a simple shear problem. *Zeitschrift für angewandte Mathematik und Physik* **59**, 904–917 (2008)
- P. Perzyna, The constitutive equations for rate sensitive plastic materials. *Q. Appl. Math.* **20**, 321–332 (1963)
- P. Perzyna, Fundamental problems in viscoplasticity. *Adv. Appl. Mech.* **9**, 243–377 (1966)
- P. Perzyna, *Termodynamika materiałów niesprężystych* (PWN, Warszawa, 1978) (in Polish)
- P. Perzyna, Internal state variable description of dynamic fracture of ductile solids. *Int. J. Solids Struct.* **22**, 797–818 (1986a)
- P. Perzyna, Constitutive modelling for brittle dynamic fracture in dissipative solids. *Arch. Mech.* **38**, 725–738 (1986b)
- P. Perzyna, Instability phenomena and adiabatic shear band localization in thermoplastic flow process. *Acta Mech.* **106**, 173–205 (1994)
- P. Perzyna, Constitutive modelling of dissipative solids for localization and fracture, in *Localization and Fracture Phenomena in Inelastic Solids*, Chapter 3. CISM Course and Lectures, vol. 386, ed. by P. Perzyna (Springer, 1998), pp. 99–241
- P. Perzyna, The thermodynamical theory of elasto-viscoplasticity. *Eng. Trans.* **53**, 235–316 (2005)
- P. Perzyna, The thermodynamical theory of elasto-viscoplasticity accounting for microshear banding and induced anisotropy effects. *Mechanics* **27**(1), 25–42 (2008)
- P. Perzyna, The thermodynamical theory of elasto-viscoplasticity for description of nanocrystalline metals. *Eng. Trans.* **58**(1–2), 15–74 (2010)
- P. Perzyna, Multiscale constitutive modelling of the influence of anisotropy effects on fracture phenomena in inelastic solids. *Eng. Trans.* **60**(3), 225–284 (2012)
- I. Podlubny, Fractional differential equations, in *Mathematics in Science and Engineering*, vol. 198 (Academic Press, USA, 1999)
- D. Polyzos, D.I. Fotiadis, Derivation of Mindlin's first and second strain gradient elastic theory via simple lattice and continuum models. *Int. J. Solids Struct.* **49**, 470–480 (2012)

- C. Rymarz, *Mechanika ośrodków* (PWN, Warszawa, 1993) (in Polish)
- L. Seaman, D.R. Curran, D.A. Shockey, Computational models for ductile and brittle fracture. *J. Appl. Phys.* **47**(11), 4814–4826 (1976)
- S. Shima, M. Oyane, Plasticity for porous solids. *Int. J. Mech. Sci.* **18**, 285–291 (1976)
- D.A. Skolnik, H.T. Liu, H.C. Wu, L.Z. Sun, Anisotropic elastoplastic and damage behavior of sic/al composite sheets. *Int. J. Damage Mech.* **17**, 247–272 (2008)
- L.J. Sluys, Wave Propagation, Localization and Dispersion in Softening Solids. Doctoral Thesis, Delft University Press, Delft, 1992
- J.-H. Song, H. Wang, T. Belytschko, A comparative study on finite element methods for dynamic fracture. *Comput. Mech.* **42**, 239–250 (2008)
- W. Sumelka, The Constitutive Model of the Anisotropy Evolution for Metals with Microstructural Defects, Publishing House of Poznan University of Technology, Poznań, 2009
- W. Sumelka, Role of covariance in continuum damage mechanics. *ASCE J. Eng. Mech.* **139**(11), 1610–1620 (2013)
- W. Sumelka, Fractional viscoplasticity. *Mech. Res. Commun.* **56**, 31–36 (2014)
- W. Sumelka, A. Glema, The evolution of microvoids in elastic solids, in *17th International Conference on Computer Methods in Mechanics CMM-2007*, Łódź-Spała, 19–22 June 2007, pp. 347–348
- W. Sumelka, A. Glema, Intrinsic microstructure anisotropy in elastic solids, in *GAMM 2008 79th Annual Meeting of the International Association of Applied Mathematics and Mechanics*, Bremen, 31 Mar–4 Apr 2008
- W. Sumelka, T. Łodygowski, The influence of the initial microdamage anisotropy on macrodamage mode during extremely fast thermomechanical processes. *Arch. Appl. Mech.* **81**(12), 1973–1992 (2011)
- W. Sumelka, T. Łodygowski, Reduction of the number of material parameters by ANN approximation. *Comput. Mech.* **52**, 287–300 (2013)
- W. Sumelka, M. Nowak, Non-normality and induced plastic anisotropy under fractional plastic flow rule: a numerical study. *Int. J. Numer. Anal. Methods Geomech.* **40**, 651–675 (2016)
- W. Sumelka, M. Nowak, On a general numerical scheme for the fractional plastic flow rule. *Mech. Mater.* (2017). <https://doi.org/10.1016/j.mechmat.2017.02.005>
- Y. Sun, Y. Shen, Constitutive model of granular soils using fractional-order plastic-flow rule. *Int. J. Geomech.* **17**(8), 04017025 (2017)
- V.E. Tarasov, General lattice model of gradient elasticity. *Mod. Phys. Lett. B* **28**(17), 1450054 (2014)
- C. Teodosiu, F. Sidoroff, A theory of finite elastoviscoplasticity of single crystals. *Int. J. Eng. Sci.* **14**(2), 165–176 (1976)
- R.A. Toupin, Elastic materials with couple-stresses. *Arch. Ration. Mech. Anal.* **11**(5), 385–414 (1963)
- R.A. Toupin, Theories of elasticity with couple-stress. *Arch. Ration. Mech. Anal.* **17**(2), 85–112 (1964)
- C. Truesdell, W. Noll, The non-linear field theories of mechanics, in *Handbuch der Physik*, vol. III/3, ed. by S. Flügge (Springer, Berlin, 1965)
- G.Z. Voyiadjis, F.H. Abed, Microstructural based models for bcc and fcc metals with temperature and strain rate dependency. *Mech. Mater.* **37**, 355–378 (2005)
- G.Z. Voyiadjis, F.H. Abed, Implicit algorithm for finite deformation hypoelastic-viscoplasticity in fcc metals. *Int. J. Numer. Methods Eng.* **67**, 933–959 (2006)
- G.Z. Voyiadjis, R.K. Abu Al-Rub, Gradient plasticity theory with a variable length scale parameter. *Int. J. Solids Struct.* **42**(14), 3998–4029 (2005)
- G.Z. Voyiadjis, P.I. Kattan, Evolution of fabric tensors in damage mechanics of solids with micro-cracks: part I – theory and fundamental concepts. *Mech. Res. Commun.* **34**, 145–154 (2007)
- H. Xiao, O.T. Bruhns, A. Meyers, Hypo-elasticity model based upon the logarithmic stress rate. *J. Elast.* **47**, 51–68 (1997)
- H. Xiao, O.T. Bruhns, A. Meyers, Strain rates and material spin. *J. Elast.* **52**, 1–41 (1998)

-
- R. Xiao, H. Sun, W. Chen, A finite deformation fractional viscoplastic model for the glass transition behavior of amorphous polymers. *Int. J. Non Linear Mech.* **93**, 7–14 (2017)
- F. Yang, A.C.M. Chong, D.C.C. Lam, P. Tong, Couple stress based strain gradient theory for elasticity. *Int. J. Solids Struct.* **39**, 2731–2743 (2002)
- S. Zaremba, Sur une forme perfectionnée de la théorie de la relaxation. *Bull. Int. Acad. Sci. Cracovie* 594–614 (1903)



Finite Element Analysis of Thermodynamically Consistent Strain Gradient Plasticity Theory and Applications

22

George Z. Voyiadjis and Yooseob Song

Contents

Introduction	782
Kinematics	785
Principle of Virtual Power: Grain Interior	787
Energetic and Dissipative Thermodynamic Microforces: Grain Interior	789
Flow Rule: Grain Interior	795
Thermodynamic Derivations of the Heat Evolution Equation	795
Principle of Virtual Power: Grain Boundary	797
Energetic and Dissipative Thermodynamic Microforces: Grain Boundary	798
Flow Rule: Grain Boundary	800
Finite Element Formulation of the Proposed SGP Model	801
Validation of the Proposed SGP Model	804
Microfree Grain Boundary	808
Energetic Gradient Hardening	808
Dissipative Gradient Strengthening	809
GND Hardening	809
Temperature-Related Parametric Study	809
Microhard Grain Boundary	813
Intermediate (Deformable) Grain Boundary	817
Generalized Structure for Modeling Polycrystals from Micro- to Nanoscale Range	825
Conclusions	834
Appendix A. Deriving the Balance Equations	835
References	837

G. Z. Voyiadjis (✉) · Y. Song
Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge,
LA, USA
e-mail: voyiadjis@eng.lsu.edu; ysong17@lsu.edu

Abstract

In this chapter, a coupled thermomechanical gradient-enhanced continuum plasticity theory containing the flow rules of the grain interior and grain boundary areas is developed within the thermodynamically consistent framework. Two-dimensional finite element implementation for the proposed gradient plasticity theory is carried out to examine the micro-mechanical and thermal characteristics of small-scale metallic volumes. The proposed model is conceptually based on the dislocation interaction mechanisms and thermal activation energy. The thermodynamic conjugate microstresses are decomposed into dissipative and energetic components; correspondingly, the dissipative and energetic length scales for both the grain interior and grain boundary are incorporated in the proposed model, and an additional length scale related to the geometrically necessary dislocation-induced strengthening is also included. Not only the partial heat dissipation caused by the fast transient time but also the distribution of temperature caused by the transition from the plastic work to the heat is included into the coupled thermomechanical model by deriving a generalized heat equation. The derived constitutive framework and two-dimensional finite element model are validated through the comparison with the experimental observations conducted on microscale thin films. The proposed enhanced model is examined by solving the simple shear problem and the square plate problem to explore the thermomechanical characteristics of small-scale metallic materials. Finally, some significant conclusions are presented.

Keywords

Strain gradient plasticity · Thermomechanical coupling · Grain boundary · 2D FEM · Validation · Size effect

Introduction

The conventional continuum plasticity model is characteristically size-independent and is not capable of capturing the size effects, in particular, when the material is subjected to the nonhomogeneous (heterogeneous) plastic deformation under the fast transient time and its size ranges from a few hundreds of nanometers to a few tens of micrometers. The evidence of such a behavior is found in many micro-mechanical experimental observations such as nano-/micro-indentation hardness (Almasri and Voyiadjis 2010; Kim and Park 1996; Lim et al. 2017; Park et al. 1996; Voyiadjis et al. 2010; Voyiadjis and Peters 2010; Voyiadjis and Zhang 2015; Zhang and Voyiadjis 2016), nano/micro-pillars (Hwang et al. 1995), torsion of micron-dimensioned metal wires (Fleck and Hutchinson 1997; Kim et al. 1994), bending of microscale single and polycrystalline thin films (Venkatraman et al. 1994), thin beams under micron-dimensioned tension/bending (Huang et al. 2006; Parilla et al. 1993), flow strength of nanocrystalline metals (Bergeman et al. 2006), and microscale reverse extrusion of copper (Zhang et al. 2018).

It is commonly accepted that the interaction between the statistically stored dislocations (SSDs) and the geometrically necessary dislocations (GNDs) gives rise to the size effect observed in micro-/nanoscale metallic volumes. The SSDs are stored by random trapping each other and increase with the plastic strain, whereas the GNDs are stored to preserve the compatibility of diverse material components and increase with the gradient of the plastic strain. As the size of material specimen decreases, the GNDs increase the resistance to deformation by acting as the blockages to the SSDs (Fleck et al. 1994). This mechanism is called glide control because the existence of GNDs, caused by nonuniform deformation or the prescribed boundary conditions, restrains the slip of dislocation glide (Muhlhaus and Aifantis 1991; Nicola et al. 2006; Xiang and Vlassak 2006). Another mechanism for the size effect is the dislocation starvation caused by the insufficient amount of dislocations that arises from the small volumes (Bergeman et al. 2006; Giacomazzi et al. 2004; Yaghoobi and Voyiadjis 2016).

Numerous theoretical and numerical works have been carried out to explore the aforementioned phenomena based on the gradient-enhanced nonlocal plasticity theory (Fleck and Hutchinson 1997; Gudmundson 2004; Gurtin and Anand 2009; Hutchinson 2012; Ivanitsky and Kadakov 1983; Lele and Anand 2008; Song and Voyiadjis 2018; Voyiadjis and Song 2017) since the pioneering investigations of Aifantis (1984, 1987) which incorporate the gradient term in the conventional flow rule. McDowell (2010) reviewed the trends in plasticity research for metals over the 25 years prior to the publication year (2010) in terms of the multiscale kinematics, the effect of material length scale, the role of grain boundaries, and so forth.

Hutchinson (2012) classified the strain gradient version of J2 flow theories into two classes: incremental theory developed by Fleck and Hutchinson and non-incremental theory developed by Gudmundson, Gurtin, and Anand (c.f. see Fleck et al. (2014, 2015); Gudmundson (2004); Gurtin and Anand (2005, 2009); Hutchinson (2012) for details). Fleck and co-workers (Fleck et al. 2014, 2015) then pointed out that the specific phenomenon, which exhibits a significant stress jump due to infinitesimal variation in the direction of plastic strain that may occur under the nonproportional loading, arises in the non-incremental theory and also discussed its physical acceptance in their work. Fleck and co-workers (Fleck et al. 2014, 2015) have shown this phenomenon with two plane strain problems, stretch-bending problem and stretch-surface passivation problem, for nonproportional loading condition. In their work, it is noted that the dissipative higher-order microforces always cause the significant stress gap or jump under the nonproportional loading conditions. Recently, Voyiadjis and Song (2017) examined the stress jump phenomenon with the stretch-surface passivation problem using the one-dimensional finite element implementation, and an extensive parametric study is also performed in order to investigate the characteristics of the stress jump phenomenon.

Another important issue in the strain gradient plasticity (SGP) theory is the thermal effect. In the nano/micro systems, the effect of temperature gradient needs to be considered for the fast transient time. If the mean free path of phonons is approximately the medium size, heat transfer is partly ballistic rather than purely diffusive. This is caused by the small depth of the zone influenced by heat or small size of

the structure and the non-equilibrium transition of thermodynamic circumstances related to reducing the response time (Tzou and Zhang 1995). Moreover, when the response time in nano-/microscale materials decreases to the thermalization time range, it results in the non-equilibrium conversion of thermodynamic states between phonons and electrons (Brorson et al. 1990; Tzou and Zhang 1995). The conventional heat equation is not capable of capturing the effect of electron-phonon interaction in this time frame; thus, the microscopic generalized heat equation has to be employed to interpret these phenomena.

Voyiadjis and co-workers (Voyiadjis and Deliktas 2009a, b; Voyiadjis and Faghihi 2012; Voyiadjis et al. 2014, 2017; Voyiadjis and Song 2017) have developed the thermodynamically consistent and coupled thermomechanical SGP models to study the characteristics of nano-/microscale metallic materials. All those works, however, have been limited to one-dimensional finite element implementation. As it is well known, there is bound to be a fundamental difference between one-dimensional finite element implementation and two-dimensional one. For example, in one-dimensional case, some special complications, e.g., the resonance between the physical scale and mesh scale, cannot be considered during the simulation. It should be noted that there is no difference between the different dimensions from the variational point of view; however, the difference exists in the description of the finite-dimensional approximation spaces. The finite element implementation in multidimension is based on the same principle of the one in the one dimension; thus the piecewise-polynomial functions of low degree with many terms are considered for accuracy. In the two-dimensional model, it is significantly more complicated since the polynomial functions have more variables and the open sets are much more varied than in the one-dimensional model. In terms of the dimensional extension, there is the simple modification from one-dimensional finite element implementation for the strain gradient plasticity model to the two-dimensional one in Voyiadjis and Song (2017). However, in that work, the effects of temperature and its gradient were not considered, but just addressed the effect of the mechanical component of the thermodynamic microforces in terms of the stress jump phenomenon. Recently, in Song and Voyiadjis (2018), the two-dimensional finite element implementation of the coupled thermomechanical strain gradient plasticity model is performed. In that work, two null boundary conditions, i.e., microscopically free and hard boundary conditions, are considered at the grain boundary to describe the dislocation movement and the plastic flow at the grain boundary areas.

It is well known that the free surface may act as the source for the defect development and its propagation toward the grain inside, whereas the grain boundaries block this dislocation movement and, consequently, give rise to the strain gradients to accommodate the geometrically necessary dislocations (Hirth and Lothe 1982). In addition, the grain boundaries can be the source of dislocation through the transmission of plastic slip to the neighboring grains (Clark et al. 1992). Besides these physical remarks, from the mathematical viewpoint, the nonstandard boundary conditions are necessary at the external boundary of a region for the well-posed governing equations in the implementation of higher-order strain gradient plasticity

models. Therefore, careful modeling of the grain boundary is important in the continued development of higher-order strain gradient plasticity models.

The experimental observations on slip transmission motivate one to assume that the effect of surface/interfacial energy and the global nonlocal energy residual should be nonvanishing. Examples can be found from the in situ TEM direct observations, e.g., Lee et al. (1989), or using the geometrically necessary dislocation (GND) concept in the description of observations in bicrystallines, e.g., Sun et al. (2000), and nanoindentation tests close to the grain boundary, e.g., Soer et al. (2005). This results in a new type of boundary condition, in the context of strain gradient plasticity incorporating the interfacial energy, accounting for the surface resistance to the slip transfer due to the grain boundary misalignment (see, e.g., Aifantis and Willis (2005); Cermelli and Gurtin (2002); Fredriksson and Gudmundson (2007); Gudmundson (2004); Gurtin (2008)).

Also, in this chapter, two-dimensional numerical simulation in the context of the small deformation framework is developed incorporating the temperature- and rate-dependent flow rules for the grain interior and grain boundary, and the proposed model is validated by comparing against two sets of small-scale experiments showing the size effects. The reason for choosing these two experiments is that these experiments were also used in Voyiadjis and Song (2017) for validating the one-dimensional model, so it will be possible to compare the numerical results and validated material properties directly. The simple shear problem is solved based on the validated model in order to examine the size effect in the small-scale metallic materials. The square plate problem with the various grain boundary conditions is also solved to investigate the grain boundary effect in conjunction with the length scales.

Additionally, a generalized structure for modeling polycrystals from micro- to nano-size range is introduced Voyiadjis and Deliktas (2010). The polycrystal structure is defined in terms of the grain core, the grain boundary, and the triple junction regions with their corresponding volume fractions. Depending on the size of the crystal from micro to nano, different types of analyses are used for the respective different regions of the polycrystal. The analyses encompass local and nonlocal continuum or crystal plasticity. Depending on the physics of the region, dislocation-based inelastic deformation and/or slip/separation is used to characterize the behavior of the material. The analyses incorporate interfacial energy with grain boundary sliding and grain boundary separation. Certain state variables are appropriately decomposed into energetic and dissipative components to accurately describe the size effects.

Kinematics

In this chapter, tensors are denoted by the subscripts i, j, k, l, m , and n . The superscripts e, p, int, ext, en, dis , etc. imply specific quantities such as elastic state, plastic state, internal, external, energetic, dissipative, etc., respectively. Also, the superimposed dot represents derivative with respect to time, and the indices after a comma represent the partial derivatives.

In the conventional continuum plastic theory of the isotropic solids for the small deformation assumption, the displacement gradient $u_{i,j}$ is decomposed into elastic $u_{i,j}^e$ and plastic counterparts $u_{i,j}^p$ as follows:

$$u_{i,j} = u_{i,j}^e + u_{i,j}^p \quad \text{where } u_{k,k}^p = 0 \quad (1)$$

where the elastic distortion $u_{i,j}^e$ and the plastic distortion $u_{i,j}^p$ indicate the recoverable stretching and rotation and the evolution of dislocations in the material structure, respectively.

For the small deformation framework in the conventional theory, the strain tensor ε_{ij} is also decomposed into the elastic and plastic elements as follows:

$$\varepsilon_{ij} = \varepsilon_{ij}^e + \varepsilon_{ij}^p = \frac{1}{2} (u_{i,j}^e + u_{j,i}^e) + \frac{1}{2} (u_{i,j}^p + u_{j,i}^p) \quad (2)$$

Here, the plastic strain is assumed to be deviatoric, $\varepsilon_{kk}^p = 0$, since the material volume is not changed by the dislocation glide-induced plastic deformation (Gurtin and Anand 2009). The plastic rotation is then given as follows:

$$W_{ij}^p = \frac{1}{2} (u_{i,j}^p - u_{j,i}^p) \quad (3)$$

In the conventional isotropic plasticity theory, since the plastic rotations (the rotation of material relative to the lattice) may be absorbed by their elastic counterparts without any influence on the field equations, they are fundamentally irrelevant to the theory. In this sense, the plasticity irrotational assumption is principally adopted in this chapter as indicated in Eq. (4) (see Gurtin et al. (2010)).

$$W_{ij}^p = 0 \quad (4)$$

Accordingly, $u_{i,j}^p = \varepsilon_{ij}^p$.

Meanwhile, the direction of plastic flow N_{ij} is given by

$$N_{ij} = \frac{\dot{\varepsilon}_{ij}^p}{\|\dot{\varepsilon}_{ij}^p\|} = \frac{\dot{\varepsilon}_{ij}^p}{\dot{e}^p} \Rightarrow \dot{\varepsilon}_{ij}^p = \dot{e}^p N_{ij} \quad (5)$$

with the accumulated plastic strain rate \dot{e}^p given by

$$\dot{e}^p = \|\dot{\varepsilon}_{ij}^p\| = \sqrt{\dot{\varepsilon}_{ij}^p \dot{\varepsilon}_{ij}^p} \quad (6)$$

Correspondingly, the accumulated plastic strain e^p can be obtained by

$$e^p = \int_0^t \dot{e}^p dt \quad (7)$$

Principle of Virtual Power: Grain Interior

Recently, the thermodynamically consistent SGP theories have been developed by Voyiadjis and his co-workers to account for the characteristics of small-scale metallic material behavior based on the principle of virtual power (Voyiadjis and Song 2017; Voyiadjis et al. 2017).

The internal power \mathcal{P}^{int} is presented with a combination of three energy contributions, i.e., the macro-, micro-, and thermal-energy contributions, in the arbitrary region Ω_0 as follows:

$$\mathcal{P}^{int} = \int_{\Omega_0} \left(\underbrace{\sigma_{ij} \dot{\epsilon}_{ij}^e}_{\text{Macro}} + \underbrace{x \dot{e}^p + Q_i \dot{e}_i^p}_{\text{Micro}} + \underbrace{\mathcal{A} \dot{\mathcal{T}} + \mathcal{B}_i \dot{\mathcal{T}}_{,i}}_{\text{Thermal}} \right) dV \quad (8)$$

where x and Q_i are the thermodynamic microforces conjugate respectively to \dot{e}^p and \dot{e}_i^p and σ_{ij} is the Cauchy stress tensor. It is assumed in this chapter that extra contributions to the internal power exist from the temperature. \mathcal{A} and \mathcal{B}_i are the micromorphic scalar and vector generalized stress-like variables conjugate to the temperature rate $\dot{\mathcal{T}}$ and the first gradient of the temperature rate $\dot{\mathcal{T}}_{,i}$, respectively. These terms are introduced in a micromorphic fashion to lead the additional thermal balance equations considering the nonlocal thermal effects (Liu et al. 2017). Therefore, the purely mechanical part of the internal power is complemented by the thermal contributions that represent the thermal part of the power of work, which is the convectively performed power. Note that only the first gradient of temperature rate is considered, for the sake of simplicity. Meanwhile, one can refer to Faghihi and Voyiadjis (2014) to see the model in the absence of the temperature-related terms \mathcal{A} and \mathcal{B}_i .

The internal power \mathcal{P}^{int} for Ω_0 is equated with the external power \mathcal{P}^{ext} expended by the macro- and microtractions $(\boldsymbol{\tau}_i, \boldsymbol{m})$ on the external surface $\partial\Omega_0$ and the body forces acting within Ω_0 as shown below:

$$\mathcal{P}^{ext} = \int_{\Omega_0} \underbrace{\boldsymbol{\ell}_i \dot{\boldsymbol{u}}_i}_{\text{Macro}} dV + \int_{\partial\Omega_0} \left(\underbrace{\boldsymbol{\tau}_i \dot{\boldsymbol{u}}_i}_{\text{Macro}} + \underbrace{\boldsymbol{m} \dot{e}^p}_{\text{Micro}} + \underbrace{\boldsymbol{a} \dot{\mathcal{T}}}_{\text{Thermal}} \right) dS \quad (9)$$

where $\boldsymbol{\ell}_i$ is the generalized external body force conjugate to the macroscopic velocity $\dot{\boldsymbol{u}}_i$. Furthermore, it is assumed for the external power to have the term of \boldsymbol{a} , conjugate to $\dot{\mathcal{T}}$ for the thermal effect.

By using the equation, $\mathcal{P}^{int} = \mathcal{P}^{ext}$, in conjunction with the divergence theorem and factoring out the common terms, the balance equations for the macroscopic linear momentum, nonlocal microforce, and generalized stresses \mathcal{A} and \mathcal{B}_i for volume Ω_0 can be obtained, respectively, as follows (see Appendix A for detailed derivations):

$$\sigma_{ij,j} + \mathcal{L}_i = 0 \quad (10)$$

$$\bar{\sigma}_{ij} = (\mathbf{x} - Q_{k,k}) N_{ij} \quad (11)$$

$$\mathcal{B}_{i,i} - \mathcal{A} = 0 \quad (12)$$

where $\bar{\sigma}_{ij}$ is the deviatoric part of σ_{ij} with the Kronecker delta δ_{ij} ($\bar{\sigma}_{ij} = \sigma_{ij} - \sigma_{kk} \delta_{ij} / 3$).

On $\partial\Omega_0$, the balance equations for the local surface traction and the nonlocal microtraction are expressed with the outward unit normal vector to $\partial\Omega_0$, n_i , respectively, as

$$\mathcal{t}_j = \sigma_{ij} n_i \quad (13)$$

$$\mathcal{m} = Q_i n_i \quad (14)$$

$$\mathcal{a} = \mathcal{B}_i n_i \quad (15)$$

The first law of thermodynamics is considered here to derive the thermodynamically consistent formulation to account for the thermo-viscoplastic small-scale behavior of the metallic volumes during the fast transient time. The adiabatic viscoplastic deformation for metals is affected by the initial temperature, the loading rate, and the temperature evolution caused by the transition from plastic work to heat. The enhanced gradient theory is employed for the mechanical part of the formulation, whereas the micromorphic model is employed for the thermal part as follows (see the work of Forest and Amestoy (2008)):

$$\rho \dot{\mathcal{E}} = \sigma_{ij} \dot{\epsilon}_{ij}^e + \mathbf{x} \dot{e}^p + Q_i \dot{e}_i^p + \mathcal{A} \dot{\mathcal{T}} + \mathcal{B}_i \dot{\mathcal{T}}_{,i} - q_{i,i} + \rho \mathcal{H}^{ext} \quad (16)$$

where ρ is the mass density, \mathcal{E} is the specific internal energy, q_i is the thermal flux vector, and \mathcal{H}^{ext} is the specific heat from the external source.

The second law of thermodynamics introduces a physical base to account for the GNDs' distribution in the body. The following entropy production inequality can be obtained based on the basic statement of this law that the free energy must increase at a rate less than the one at which the work is carried out, with the specific entropy s and the micromorphic approach by Forest (2009):

$$-\rho \dot{\mathcal{E}} + \rho \dot{s} \mathcal{T} + \sigma_{ij} \dot{\epsilon}_{ij}^e + \mathbf{x} \dot{e}^p + Q_i \dot{e}_i^p + \mathcal{A} \dot{\mathcal{T}} + \mathcal{B}_i \dot{\mathcal{T}}_{,i} - q_i \frac{\mathcal{T}_{,i}}{\mathcal{T}} \geq 0 \quad (17)$$

The entropy production vector is assumed in this chapter to be equal to the thermal flux vector divided by the temperature, as given in Coleman and Noll (1963).

Energetic and Dissipative Thermodynamic Microforces: Grain Interior

The Helmholtz free energy Ψ (per unit volume) is obtained with the entropy \mathcal{s} , internal energy \mathcal{E} , and temperature \mathcal{T} describing a current state of the material as follows:

$$\Psi = \mathcal{E} - \mathcal{T}\mathcal{s} \quad (18)$$

By using Eqs. (17) and (18), the Clausius-Duhem inequality is derived as follows:

$$\sigma_{ij}\dot{\varepsilon}_{ij}^e + \mathbf{x}\dot{e}^p + Q_i\dot{e}_{,i}^p + \mathcal{A}\dot{\mathcal{T}} + \mathcal{B}_i\dot{\mathcal{T}}_{,i} - \rho\dot{\Psi} - \rho\mathcal{s}\dot{\mathcal{T}} - q_i\frac{\mathcal{T}_{,i}}{\mathcal{T}} \geq 0 \quad (19)$$

For deriving the constitutive equations within a small-scale framework, an attempt to address the effect of the nonuniform microdefect distribution with the temperature is carried out in the present work with the functional form of the Helmholtz free energy given as:

$$\Psi = \Psi\left(\varepsilon_{ij}^e, e^p, e_{,i}^p, \mathcal{T}, \mathcal{T}_{,i}\right) \quad (20)$$

In the process of developing the constitutive equations, the plastic dissipation work must be nonnegative. By taking the time derivative of the Helmholtz free energy, $\dot{\Psi}$ is expressed as follows:

$$\dot{\Psi} = \frac{\partial\Psi}{\partial\varepsilon_{ij}^e}\dot{\varepsilon}_{ij}^e + \frac{\partial\Psi}{\partial e^p}\dot{e}^p + \frac{\partial\Psi}{\partial e_{,i}^p}\dot{e}_{,i}^p + \frac{\partial\Psi}{\partial\mathcal{T}}\dot{\mathcal{T}} + \frac{\partial\Psi}{\partial\mathcal{T}_{,i}}\dot{\mathcal{T}}_{,i} \quad (21)$$

Substituting Eq. (21) into Eq. (19) and factoring the common terms out give the inequality as follows:

$$\begin{aligned} &\left(\sigma_{ij} - \rho\frac{\partial\Psi}{\partial\varepsilon_{ij}^e}\right)\dot{\varepsilon}_{ij}^e + \left(\mathbf{x} - \rho\frac{\partial\Psi}{\partial e^p}\right)\dot{e}^p + \left(Q_i - \rho\frac{\partial\Psi}{\partial e_{,i}^p}\right)\dot{e}_{,i}^p + \left(\mathcal{A} - \rho\mathcal{s} - \rho\frac{\partial\Psi}{\partial\mathcal{T}}\right)\dot{\mathcal{T}} \\ &+ \left(\mathcal{B}_i - \rho\frac{\partial\Psi}{\partial\mathcal{T}_{,i}}\right)\dot{\mathcal{T}}_{,i} - \frac{q_i}{\mathcal{T}}\mathcal{T}_{,i} \geq 0 \end{aligned} \quad (22)$$

Meanwhile, the thermodynamic conjugate microforces \mathbf{x} , Q_i , and \mathcal{A} are assumed to be decomposed into the energetic and the dissipative elements as follows:

$$\mathbf{x} = \mathbf{x}^{en} + \mathbf{x}^{dis} \quad (23)$$

$$Q_i = Q_i^{en} + Q_i^{dis} \quad (24)$$

$$\mathcal{A} = \mathcal{A}^{en} + \mathcal{A}^{dis} \quad (25)$$

Substituting Eqs. (23), (24), and (25) into Eq. (22) and rearranging them in accordance with the energetic and the dissipative parts result in the following expression:

$$\begin{aligned} & \left(\sigma_{ij} - \rho \frac{\partial \Psi}{\partial \varepsilon_{ij}^e} \right) \dot{\varepsilon}_{ij}^e + \left(x^{en} - \rho \frac{\partial \Psi}{\partial e^p} \right) \dot{e}^p + \left(Q_i^{en} - \rho \frac{\partial \Psi}{\partial e_{,i}^p} \right) \dot{e}_{,i}^p \\ & + \left(\mathcal{A}^{en} - \rho s - \rho \frac{\partial \Psi}{\partial \mathcal{T}} \right) \dot{\mathcal{T}} + \left(B_i - \rho \frac{\partial \Psi}{\partial \mathcal{T}_{,i}} \right) \dot{\mathcal{T}}_{,i} + x^{dis} \dot{e}^p \\ & + Q_i^{dis} \dot{e}_{,i}^p + \mathcal{A}^{dis} \dot{\mathcal{T}} - \frac{q_i}{\mathcal{T}} \mathcal{T}_{,i} \geq 0 \end{aligned} \quad (26)$$

By assuming that the fifth term in Eq. (26) is strictly energetic, the energetic components of the thermodynamic microforces are defined as follows:

$$\sigma_{ij} = \rho \frac{\partial \Psi}{\partial \varepsilon_{ij}^e} \quad (27)$$

$$x^{en} = \rho \frac{\partial \Psi}{\partial e^p} \quad (28)$$

$$Q_i^{en} = \rho \frac{\partial \Psi}{\partial e_{,i}^p} \quad (29)$$

$$\mathcal{A}^{en} = \rho \left(s + \frac{\partial \Psi}{\partial \mathcal{T}} \right) \quad (30)$$

$$B_i = \rho \frac{\partial \Psi}{\partial \mathcal{T}_{,i}} \quad (31)$$

The dissipation density per unit time \mathcal{D} is then obtained as:

$$\mathcal{D} = x^{dis} \dot{e}^p + Q_i^{dis} \dot{e}_{,i}^p + \mathcal{A}^{dis} \dot{\mathcal{T}} - \frac{q_i}{\mathcal{T}} \mathcal{T}_{,i} \geq 0 \quad (32)$$

The dissipative counterparts of the thermodynamic microforces are obtained from the dissipation potential $\mathcal{D}(\dot{e}^p, \dot{e}_{,i}^p, \dot{\mathcal{T}}, \mathcal{T}_{,i})$ as follows:

$$x^{dis} = \frac{\partial \mathcal{D}}{\partial \dot{e}^p} \quad (33)$$

$$Q_i^{dis} = \frac{\partial \mathcal{D}}{\partial \dot{e}_{,i}^p} \quad (34)$$

$$\mathcal{A}^{dis} = \frac{\partial \mathcal{D}}{\partial \dot{\mathcal{T}}} \quad (35)$$

$$-\frac{q_i}{\mathcal{T}} = \frac{\partial \mathcal{D}}{\partial \mathcal{T}_{,i}} \quad (36)$$

It is necessary to define the proper formulation of Helmholtz free energy Ψ because it establishes the basis for the derivation of the constitutive relations. It has to take not only the material type such as fcc metals, bcc metals, polymer, steel alloys, concrete, etc. but also the deformation condition such as rate dependency of the material into consideration. In the current work, the Helmholtz free energy function is put forward with three main counterparts, i.e., elastic energy Ψ^e , defect energy Ψ^d , and thermal energy Ψ^{th} , as follows (Voyiadjis and Song 2017; Voyiadjis et al. 2017):

$$\Psi \left(\varepsilon_{ij}^e, e^p, e_{,i}^p, \mathcal{T}, \mathcal{T}_{,i} \right) = \Psi^e \left(\varepsilon_{ij}^e, \mathcal{T} \right) + \Psi^d \left(e^p, e_{,i}^p, \mathcal{T} \right) + \Psi^{th} \left(\mathcal{T}, \mathcal{T}_{,i} \right) \quad (37)$$

with

$$\Psi^e \left(\varepsilon_{ij}^e, \mathcal{T} \right) = \frac{1}{2\rho} \varepsilon_{ij}^e E_{ijkl} \varepsilon_{kl}^e - \frac{\alpha^{th}}{\rho} (\mathcal{T} - \mathcal{T}_r) \varepsilon_{ij}^e \delta_{ij} \quad (38)$$

$$\Psi^d \left(e^p, e_{,i}^p, \mathcal{T} \right) = \underbrace{\frac{\mathcal{H}_0}{\rho(\mathcal{r} + 1) \left[1 - \left(\frac{\mathcal{T}}{\mathcal{T}_y} \right)^n \right]} (e^p)^{\mathcal{r}+1}}_{\Psi_1^d} + \underbrace{\frac{\sigma_0}{\rho(\vartheta + 1)} [\ell_{en}^2 (e_{,i}^p e_{,i}^p)]^{\frac{\vartheta+1}{2}}}_{\Psi_2^d} \quad (39)$$

$$\Psi^{th} \left(\mathcal{T}, \mathcal{T}_{,i} \right) = -\frac{1}{2} \frac{c_\varepsilon}{\mathcal{T}_r} (\mathcal{T} - \mathcal{T}_r)^2 - \frac{1}{2\rho} a \mathcal{T}_{,i} \mathcal{T}_{,i} \quad (40)$$

where α^{th} is the thermal expansion coefficient, E_{ijkl} is the elastic modulus tensor, \mathcal{H}_0 is the standard isotropic hardening parameter, \mathcal{r} ($0 < \mathcal{r} < 1$) is the isotropic hardening material parameter, \mathcal{T}_y and n are the thermal material parameters, $\sigma_0 > 0$ is the stress-dimensioned scaling parameter to explain the initial slip resistance, ℓ_{en} is the energetic material length scale describing the feature of the short-range interaction of the GNDs, a is the material constant for the isotropic heat conduction which accounts for the interaction of the energy carriers, ϑ is the parameter for governing the nonlinearity of the gradient-dependent defect energy, $\mathcal{T}_r > 0$ is the reference temperature, and c_ε is the specific heat capacity at the constant stress. In this chapter, \mathcal{T}_y is determined by the calibration with experimental data.

The second term of the defect energy Ψ_2^d is postulated as a function of $e_{,i}^p$, and the condition $\vartheta > 0$ ensures the convexity of Ψ_2^d . (The short-range interaction of the coupling dislocations shifting on the adjacent slip planes, so-called GND core energy, is characterized by the defect energy Ψ_2^d . Ψ_2^d may be assumed to be a function of the GND density. It is assumed in this work that the plastic strain gradient is viewed as a macroscopic measure of GNDs.) In particular, by setting $\vartheta = 1$, the function Ψ_2^d turns to the quadratic formula (see Bardella (2006); Gurtin

(2004)), while the L_2 -norm of e_i^p is used in Garroni et al. (2010) and Ohno and Okumura (2007) by setting $\vartheta = 0$. It is worth mentioning that Ψ_2^d is independent of the temperature since it indicates the energy carried by the dislocations; thus it is energetic in nature (Lele and Anand 2008).

Since the central objective of the present work is to account for thermal variation, thermal terms are included in the free energy. The elastic and defect parts of the Helmholtz free energy functions, Ψ^e and Ψ^d , are locally convex with respect to strain-related terms at all points of the body in the equilibrium state. However, Ψ^t is a concave function of the temperature (Lubarda 2008). The second-order variation of Ψ is related to the second-order variation of \mathcal{E} by using Eq. (18) and taking the virtual variations of temperature and entropy states into account such as $\partial^2\Psi/\partial\mathcal{T}^2(\delta\mathcal{T})^2 = -\partial^2\mathcal{E}/\partial s^2(\delta s)^2$. Thus, \mathcal{E} is convex with respect to entropy s , and Ψ is concave with respect to temperature since $\partial^2\mathcal{E}/\partial s^2(\delta s)^2 > 0 \rightarrow \partial^2\Psi/\partial\mathcal{T}^2(\delta\mathcal{T})^2 < 0$. In addition, it can be assumed that entropy increases monotonically with respect to temperature; thus $\partial s/\partial\mathcal{T} = -\partial^2\Psi/\partial\mathcal{T}^2$ (Callen 1985).

One can now obtain the energetic thermodynamic forces by using the definitions in Eqs. (27), (28), (29), (30), and (31) in conjunction with Eqs. (37), (38), (39), and (40) as follows:

$$\sigma_{ij} = E_{ijkl}\varepsilon_{kl}^e - \alpha^{th}(\mathcal{T} - \mathcal{T}_r)\delta_{ij} \quad (41)$$

$$x^{en} = \mathcal{H}_0 \left[1 - \left(\frac{\mathcal{T}}{\mathcal{T}_y} \right)^n \right] (e^p)^r \quad (42)$$

$$Q_i^{en} = \sigma_0 \ell_{en}^2 \left[\ell_{en}^2 (e_{,k}^p e_{,k}^p) \right]^{\frac{\vartheta-1}{2}} e_{,i}^p \quad (43)$$

$$\mathcal{A}^{en} = \rho s - \alpha^{th}(\mathcal{T} - \mathcal{T}_r)\varepsilon_{ij}^e \delta_{ij} - \frac{c_\varepsilon}{\mathcal{T}_r}(\mathcal{T} - \mathcal{T}_r) - \frac{\mathcal{H}_0(e^p)^{r+1}}{r+1} \frac{\mathcal{T}}{\mathcal{T}_y} \left(\frac{\mathcal{T}}{\mathcal{T}_y} \right)^{n-1} \quad (44)$$

$$B_i = -a\mathcal{T}_{,i} \quad (45)$$

It is assumed here that the dissipation potential function is composed of two parts, the mechanical part which is dependent on the plastic strain and plastic strain gradient and the thermal counterpart which shows the purely thermal effect such as the heat conduction. In this sense, and in the context of Eq. (32), the functional form of the dissipation potential, which is dependent on $e_{,i}^p$, can be put forward as:

$$\mathcal{D} = \mathcal{D}^p(e^p, \varepsilon^p, \dot{e}^p, \mathcal{T}) + \mathcal{D}^g(\dot{e}_{,i}^p, \mathcal{T}) + \mathcal{D}^{th}(\dot{\mathcal{T}}, \mathcal{T}_{,i}) \quad (46)$$

where \mathcal{D}^p and \mathcal{D}^g are the mechanical parts and \mathcal{D}^{th} accounts for the purely thermal effect. The functional forms of each part are given as follows:

$$\mathcal{D}^p(e^p, \varepsilon^p, \dot{e}^p, \mathcal{T}) = \sigma_0 \sqrt{\mathcal{H}^2(e^p) + \xi^2 \left(\frac{\mu}{\sigma_0}\right)^2 b \varepsilon^p} \left[1 - \left(\frac{\mathcal{T}}{\mathcal{T}_y}\right)^n\right] \left(\frac{\dot{e}^p}{\dot{\mathcal{p}}_1}\right)^{m_1} \dot{e}^p \quad (47)$$

$$\mathcal{D}^s(\dot{e}_{,i}^p, \mathcal{T}) = \sigma_0 \left[1 - \left(\frac{\mathcal{T}}{\mathcal{T}_y}\right)^n\right] \left(\frac{\dot{\mathcal{p}}}{\dot{\mathcal{p}}_2}\right)^{m_2} \dot{\mathcal{p}} \quad (48)$$

$$\mathcal{D}^{th}(\dot{\mathcal{J}}, \mathcal{T}_{,i}) = -\frac{\zeta}{2} \dot{\mathcal{J}}^2 - \frac{1}{2} \frac{\mathcal{k}(\mathcal{T})}{\mathcal{T}} \mathcal{T}_{,i} \mathcal{T}_{,i} \quad (49)$$

where ξ and b are the numerical parameter and magnitude of the Burgers vector, which are characteristically given as $0.2 \leq \xi \leq 0.5$ and $b \approx 0.3$ nm for metals, respectively. The parameter μ is the shear modulus, $\dot{\mathcal{p}}_1$ and $\dot{\mathcal{p}}_2$ are the nonnegative reference rates, m_1 and m_2 are the nonnegative strain rate sensitivity parameters, ζ is the material constant characterizing the energy exchange between phonon and electron, and $\mathcal{k}(\mathcal{T})$ is the thermal conductivity coefficient. The dimensionless function $(\dot{e}^p/\dot{\mathcal{p}}_1)^{m_1}$ in Eq. (47) leads to different physical effects from the term $(\dot{\mathcal{p}}/\dot{\mathcal{p}}_2)^{m_2}$ in Eq. (48), in spite of the similar forms (see Lele and Anand (2008) for more details).

The parameter $\dot{\mathcal{p}}$ is a scalar measuring the plastic strain rate gradient, which is defined by

$$\dot{\mathcal{p}} \stackrel{\text{def}}{=} \ell_{dis} \|\dot{e}_{,i}^p\| = \ell_{dis} \sqrt{\dot{e}_{,i}^p \dot{e}_{,i}^p} \quad (50)$$

where ℓ_{dis} is the dissipative length scale.

The dimensionless function $\mathcal{H}(e^p)$ is related to the strain hardening/softening behavior with $\mathcal{H}(0) = 1$. In the current work, the following form of mixed hardening function is adopted (Voce 1955):

$$\mathcal{H}(e^p) = 1 + (\chi - 1) [1 - \exp(-\omega e^p)] + \frac{\mathcal{H}_0}{\sigma_0} e^p \quad (51)$$

where χ and ω are the material parameters. The strain hardening, strain softening, and strain hardening/softening can be modeled based on the particular choices for these parameters.

The Nye dislocation density tensor α_{ij} , which indicates the i -component of the resultant Burgers vector related to GNDs of line vector j , is exploited here to account for the effects of plastic strain gradient (Arsenlis and Parks 1999; Fleck and Hutchinson 1997). Nonvanishing α_{ij} indicates that the GNDs exist and the net Burgers vector b_i can be obtained by using the Stokes' theorem as follows:

$$b_i = \oint_c u_{i,k}^p dx_k = \int_s \epsilon_{jkl} u_{i,lk}^p n_j dS \quad (52)$$

where ϵ_{jkl} is the permutation tensor and n_j is the unit vector normal to the surface \mathcal{S} whose boundary is the curve \mathcal{C} . Under the assumption that the plastic flow is irrotational, the Nye dislocation density tensor α_{ij} is given by

$$\alpha_{ij} = \epsilon_{ikl} u_{j,lk}^p = \epsilon_{ikl} \epsilon_{jl,k}^p \tag{53}$$

in which, in the work of Gurtin (2004), α_{ji} is indicated as the Burgers tensor.

With neglecting the interaction between different slip systems, the total accumulation of GNDs is calculated as follows:

$$\varepsilon^p \stackrel{\text{def}}{=} \|\alpha_{ij}\| = b \rho_G \tag{54}$$

where ε^p is a scalar measure of an effective plastic strain gradient and ρ_G is the total GND density. In order to present the microstructural hardening induced by GNDs, another length scale parameter, designated as the N-G length scale parameter, is defined here:

$$\ell_{N-G} \stackrel{\text{def}}{=} \xi^2 \left(\frac{\mu}{\sigma_0} \right)^2 b \tag{55}$$

where the N-G length scale parameter was first introduced by Nix and Gao (1998). With the definition of ℓ_{N-G} , Eq. (47) can be expressed by

$$\mathcal{D}^p(e^p, \varepsilon^p, \dot{e}^p, \mathcal{T}) = \sigma_0 \sqrt{\mathcal{H}^2(e^p) + \ell_{N-G} \varepsilon^p} \left[1 - \left(\frac{\mathcal{T}}{\mathcal{T}_y} \right)^n \right] \left(\frac{\dot{e}^p}{\dot{\rho}_1} \right)^{m_1} \dot{e}^p \tag{56}$$

In the special case $\ell_{N-G} = 0$ and $\mathcal{H}(e^p) = 1$, Eq. (56) reduces to $\mathcal{D}^p = \sigma_0 \left(1 - (\mathcal{T}/\mathcal{T}_y)^n \right) (\dot{e}^p/\dot{\rho}_1)^{m_1} \dot{e}^p$, a form in Voyiadjis and Song (2017).

Using the dissipative potential given in Eqs. (47), (48), (49), and (56) along with Eqs. (33), (34), (35), and (36) and considering $\kappa(\mathcal{T})/\mathcal{T} = \kappa_0 = \text{constant}$, the constitutive relations for the dissipative microforces are obtained as follows:

$$x^{dis} = \sigma_0 \sqrt{\mathcal{H}^2(e^p) + \ell_{N-G} \varepsilon^p} \left[1 - \left(\frac{\mathcal{T}}{\mathcal{T}_y} \right)^n \right] \left(\frac{\dot{e}^p}{\dot{\rho}_1} \right)^{m_1} \tag{57}$$

$$Q_i^{dis} = \sigma_0 \ell_{dis}^2 (m_2 + 1) \left[1 - \left(\frac{\mathcal{T}}{\mathcal{T}_y} \right)^n \right] \left(\frac{\dot{\rho}}{\dot{\rho}_2} \right)^{m_2} \frac{\dot{e}_{,i}^p}{\dot{\rho}} \tag{58}$$

$$\mathcal{A}^{dis} = -\zeta \dot{\mathcal{J}} \tag{59}$$

$$\frac{q_i}{\mathcal{T}} = \kappa_0 \mathcal{T}_{,i} \tag{60}$$

Flow Rule: Grain Interior

The flow rule in the present framework is established based on the nonlocal microforce balance, given in Eq. (11), and strengthened by thermodynamically consistent constitutive relations for both energetic and dissipative microforces. By considering the backstress, the microforce equilibrium can be expressed as follows:

$$\bar{\sigma}_{ij} - \underbrace{(-Q_{k,k}^{en})}_{\text{Backstress}} N_{ij} = (x - Q_{k,k}^{dis}) N_{ij} \quad (61)$$

By substituting Eqs. (42), (43), (57), and (58) into Eq. (61), one can obtain a second-order partial differential flow rule as follows:

$$\begin{aligned} \bar{\sigma}_{ij} &= \left[-\sigma_0 \ell_{en}^2 \{ \ell_{en}^2 (e_i^p e_i^p) \}^{\frac{\theta-1}{2}} e_{,kk}^p \right] N_{ij} \\ &= \left[\mathcal{H}_0 \left[1 - \left(\frac{\mathcal{T}}{\mathcal{T}_y} \right)^n \right] (e^p)^r + \sigma_0 \sqrt{\mathcal{H}^2(e^p) + \ell_{N-G} \varepsilon^p} \right. \\ &\quad \left[1 - \left(\frac{\mathcal{T}}{\mathcal{T}_y} \right)^n \right] \left(\frac{\dot{e}^p}{\dot{p}_1} \right)^{m_1} - \sigma_0 \ell_{dis}^2 (m_2 + 1) \\ &\quad \left. \left[1 - \left(\frac{\mathcal{T}}{\mathcal{T}_y} \right)^n \right] \left(\frac{\dot{p}}{\dot{p}_2} \right)^{m_2} \frac{\dot{e}_{,kk}^p}{\dot{p}} \right] N_{ij} \end{aligned} \quad (62)$$

It is required to accompany the initial conditions for e^p and ε^p in the flow rule. A standard initial condition, for the behavior starting at time $t = 0$ from a virgin state, is assumed here such that

$$e_{t=0}^p = \varepsilon_{t=0}^p = 0 \quad (63)$$

Thermodynamic Derivations of the Heat Evolution Equation

Heat flow is controlled by the first law of thermodynamics, i.e., the energy conservation law, given in Eq. (16). The temperature field is governed by the heat flow generated through the inelastic dissipation and thermomechanical coupling effect. By considering the law of energy conservation given in Eq. (16) along with the dissipation potential given in Eqs. (32), (46), (47), (48), and (49) in conjunction with the equations for the energetic and dissipative components of the thermodynamic microforces given, respectively, by Eqs. (27), (28), (29), (30), and (31) and Eqs. (33), (34), (35), and (36), the relationship for the evolution of the entropy, which describes the irreversible process, can be derived as follows:

$$\rho \dot{\delta} \mathcal{T} = \mathcal{D} + \rho \mathcal{H}^{exi} \quad (64)$$

By using Eq. (44) for solving the rate of the entropy \dot{s} and assuming the specific heat capacity at the constant volume c_0 as $c_0 = constant \cong c_\varepsilon \mathcal{T} / \mathcal{T}_r$, the temperature evolution can be obtained as follows:

$$\begin{aligned} \rho c_0 \dot{\mathcal{T}} = & \underbrace{x^{dis} \dot{e}^p + Q_i^{dis} \dot{e}_{,i}^p}_{\text{irreversible mechanical process}} - \underbrace{\alpha^{th} (\mathcal{T} - \mathcal{T}_r) \dot{\varepsilon}_{ij}^e \delta_{ij} \mathcal{T}}_{\text{elastic-thermal coupling}} \\ & - \underbrace{\dot{\mathcal{P}} \mathcal{T}}_{\text{plastic-thermal coupling}} + \underbrace{\frac{\kappa_0}{2} \mathcal{T}_{,i} \mathcal{T}_{,i}}_{\text{heat conduction}} + \underbrace{\rho \mathcal{H}^{ext}}_{\text{heat source}} \end{aligned} \tag{65}$$

where

$$\dot{\mathcal{P}} = \mathcal{H}_0 (e^p)^r \left(\frac{\mathcal{T}}{\mathcal{T}_y} \right)^n \dot{e}^p + \frac{n \mathcal{H}_0 (e^p)^{r+1}}{(r+1) \mathcal{T}_y} \left(\frac{\mathcal{T}}{\mathcal{T}_y} \right)^{n-1} \dot{\mathcal{T}} \tag{66}$$

It should be noted that the heat conduction term can be generalized to the microscale heat equation by considering the effects of the temperature gradient on the stored energy and the temperature on the energy dissipation individually in terms of the two extra material intrinsic time scale parameters (Voyiadjis and Faghihi 2012).

By substituting the constitutive equations of the dissipative microforces into Eq. (65) and assuming that the external heat source is absent, the temperature evolution is consequently obtained as:

$$\begin{aligned} & \left[1 + \frac{n \mathcal{H}_0 (e^p)^{r+1}}{\rho c_0 (r+1) \mathcal{T}_y} \left(\frac{\mathcal{T}}{\mathcal{T}_y} \right)^{n-1} \right] \dot{\mathcal{T}} \\ & = \frac{1}{\rho c_0} \left[\left\{ \sigma_0 \sqrt{\mathcal{H}^2 (e^p) + \ell_{N-G} \varepsilon^p} \left\{ 1 - \left(\frac{\mathcal{T}}{\mathcal{T}_y} \right)^n \right\} \left(\frac{\dot{e}^p}{\dot{\mathcal{P}}_1} \right)^{m_1} \right. \right. \\ & \quad \left. \left. - \mathcal{H}_0 (e^p)^r \mathcal{T} \left(\frac{\mathcal{T}}{\mathcal{T}_y} \right)^n \right\} \dot{e}^p \right. \\ & \quad \left. + \sigma_0 \ell_{dis}^2 (m_2 + 1) \left\{ 1 - \left(\frac{\mathcal{T}}{\mathcal{T}_y} \right)^n \right\} \left(\frac{\dot{\mathcal{P}}}{\dot{\mathcal{P}}_2} \right)^{m_2} \frac{\dot{e}_{,i}^p \dot{e}_{,i}^p}{\dot{\mathcal{P}}} \right. \\ & \quad \left. - \alpha^{th} (\mathcal{T} - \mathcal{T}_r) \dot{\varepsilon}_{ij}^e \delta_{ij} \mathcal{T} \right] + \mathcal{t}_{eff} \mathcal{T}_{,ii} \end{aligned} \tag{67}$$

where the additional term \mathcal{t}_{eff} is defined as $\mathcal{t}_{eff} = \kappa_0 / 2 \rho c_0$.

Principle of Virtual Power: Grain Boundary

One of the main goals in this study is to develop the thermodynamically consistent gradient-enhanced plasticity model for the grain boundary, which should be also consistent with the one for the grain interior. Hereafter, the superscript GB and the expression GB will be used to denote the specific variables at the grain boundary.

Two grains \mathcal{G}_1 and \mathcal{G}_2 separated by the grain boundary are taken into account in this chapter, and the displacement field is assumed to be continuous, i.e., $u_i^{\mathcal{G}_1} = u_i^{\mathcal{G}_2}$, across the grain boundary. The internal part of the principle of virtual power for the grain boundary is assumed to depend on the GB accumulated plastic strain rates $\dot{\epsilon}^{p^{GB\mathcal{G}_1}}$ at $S^{GB\mathcal{G}_1}$ and $\dot{\epsilon}^{p^{GB\mathcal{G}_2}}$ at $S^{GB\mathcal{G}_2}$ in the arbitrary surface S^{GB} of the grain boundary as follows:

$$\mathcal{P}^{int GB} = \int_{S^{GB}} \left(\mathbb{M}^{GB\mathcal{G}_1} \dot{\epsilon}^{p^{GB\mathcal{G}_1}} + \mathbb{M}^{GB\mathcal{G}_2} \dot{\epsilon}^{p^{GB\mathcal{G}_2}} \right) dS^{GB} \quad (68)$$

where the GB microscopic moment tractions $\mathbb{M}^{GB\mathcal{G}_1}$ and $\mathbb{M}^{GB\mathcal{G}_2}$ are assumed to expend power over $\dot{\epsilon}^{p^{GB\mathcal{G}_1}}$ and $\dot{\epsilon}^{p^{GB\mathcal{G}_2}}$, respectively. In addition, the GB external power $\mathcal{P}^{ext GB}$ is expended by the macrotractions $\sigma_{ij}^{\mathcal{G}_1} (-n_j^{GB})$ and $\sigma_{ij}^{\mathcal{G}_2} (n_j^{GB})$ conjugate to the macroscopic velocity \dot{u}_i and the microtractions $\mathbb{Q}_k^{\mathcal{G}_1} (-n_k^{GB})$ and $\mathbb{Q}_k^{\mathcal{G}_2} (n_k^{GB})$ that are conjugate to $\dot{\epsilon}^{p^{GB\mathcal{G}_1}}$ and $\dot{\epsilon}^{p^{GB\mathcal{G}_2}}$, respectively, as follows:

$$\mathcal{P}^{ext GB} = \int_{S^{GB}} \left\{ \left(\sigma_{ij}^{\mathcal{G}_2} n_j^{GB} - \sigma_{ij}^{\mathcal{G}_1} n_j^{GB} \right) \dot{u}_i + \mathbb{Q}_k^{\mathcal{G}_2} n_k^{GB} \dot{\epsilon}^{p^{GB\mathcal{G}_2}} - \mathbb{Q}_k^{\mathcal{G}_1} n_k^{GB} \dot{\epsilon}^{p^{GB\mathcal{G}_1}} \right\} dS^{GB} \quad (69)$$

where \mathbf{n}^{GB} is the unit outward normal vector of the grain boundary surface. From $\mathcal{P}^{int GB} = \mathcal{P}^{ext GB}$, the macro- and microforce balances for the grain boundary are obtained as follows:

$$\left(\sigma_{ij}^{\mathcal{G}_1} - \sigma_{ij}^{\mathcal{G}_2} \right) n_j^{GB}; \quad \mathbb{M}^{GB\mathcal{G}_1} + \mathbb{Q}_k^{\mathcal{G}_1} n_k^{GB} = 0; \quad \mathbb{M}^{GB\mathcal{G}_2} - \mathbb{Q}_k^{\mathcal{G}_2} n_k^{GB} = 0 \quad (70)$$

The first and second laws of thermodynamics are considered to construct the thermodynamically consistent gradient- and temperature-enhanced framework for the grain boundary as follows:

$$\dot{\mathcal{E}}^{GB} = \mathbb{M}^{GB} \dot{\epsilon}^{p^{GB}} + q_i^{GB} n_i^{GB} \quad (71)$$

$$\dot{\mathcal{S}}^{GB} \mathcal{T}^{GB} - q_i^{GB} n_i^{GB} \geq 0 \quad (72)$$

where \mathcal{E}^{GB} is the GB surface energy density, q_i^{GB} is the GB heat flux vector, and \mathcal{s}^{GB} is the surface density of the entropy of the grain boundary.

Energetic and Dissipative Thermodynamic Microforces: Grain Boundary

By using the time derivative of the equation, $\Psi^{GB} = \mathcal{E}^{GB} - \mathcal{T}^{GB} \mathcal{s}^{GB}$, and substituting it into Eqs. (71) and (72), the following Clausius-Duhem inequality for the grain boundary is obtained:

$$\mathbb{M}^{GB} \dot{e}^{p^{GB}} - \dot{\Psi}^{GB} - \mathcal{s}^{GB} \dot{\mathcal{T}}^{GB} \geq 0 \quad (73)$$

One assumes the isothermal condition for the grain boundary ($\dot{\mathcal{T}}^{GB} = 0$) and the Helmholtz free energy for the grain boundary is given by $\Psi^{GB} = \Psi^{GB}(e^{p^{GB}})$. Substituting the time derivative of Ψ^{GB} into Eq. (73) gives the following inequality:

$$\mathbb{M}^{GB} \dot{e}^{p^{GB}} - \rho \frac{\partial \Psi^{GB}}{\partial e^{p^{GB}}} \dot{e}^{p^{GB}} \geq 0 \quad (74)$$

The GB thermodynamic microforce quantity \mathbb{M}^{GB} is further assumed to be decomposed into the energy and dissipative components such as $\mathbb{M}^{GB} = \mathbb{M}^{GB,en} + \mathbb{M}^{GB,dis}$. The components $\mathbb{M}^{GB,en}$ and $\mathbb{M}^{GB,dis}$ indicate the mechanisms for the pre- and post-slip transfer and thus involve the plastic strain at the grain boundary prior to the slip transfer $e^{p^{GB}(pre)}$ and the one after the slip transfer $e^{p^{GB}(post)}$, respectively, ($e^{p^{GB}} = e^{p^{GB}(pre)} + e^{p^{GB}(post)}$). From Eq. (74) one obtains

$$\left(\mathbb{M}^{GB,en} - \rho \frac{\partial \Psi^{GB}}{\partial e^{p^{GB}}} \right) \dot{e}^{p^{GB}} + \mathbb{M}^{GB,dis} \dot{e}^{p^{GB}} \geq 0 \quad (75)$$

The GB energetic microforce can be obtained as:

$$\mathbb{M}^{GB,en} = \rho \frac{\partial \Psi^{GB}}{\partial e^{p^{GB}}} \quad (76)$$

Hence the GB dissipative microforce can then be obtained as:

$$\mathbb{M}^{GB,dis} = \frac{\partial \mathcal{D}^{GB}}{\partial \dot{e}^{p^{GB}}} \quad (77)$$

where \mathcal{D}^{GB} is the nonnegative dissipation density per unit time for the grain boundary, given by $\mathcal{D}^{GB} = \mathbb{M}^{GB,dis} \dot{e}^{p^{GB}} \geq 0$. This nonnegative plastic dissipation condition can be satisfied when the GB plastic dissipation potential is a convex function of the GB accumulated plastic strain rate.

In this chapter, it is assumed by following Fredriksson and Gudmundson (2007) that the GB Helmholtz free energy per unit surface has the form of the general power law as follows:

$$\Psi^{GB} \left(e^{pGB} \right) = \frac{1}{2} G \ell_{en}^{GB} \left(e^{pGB(pre)} \right)^2 \quad (78)$$

where G is the shear modulus in the case of isotropic linear elasticity and ℓ_{en}^{GB} is the GB energetic length scale. By substituting Eq. (78) into Eq. (76), the GB energetic microforce quantity can be obtained as follows:

$$\mathbb{M}^{GB,en} = G \ell_{en}^{GB} e^{pGB(pre)} \quad (79)$$

Note that $\mathbb{M}^{GB,en}$ is independent of the plastic strain rate and temperature since this variable comes from the recoverable stored energy.

Meanwhile, two major factors might be identified affecting the energy dissipation when the dislocations move in the grain boundary area (Aifantis and Willis 2005). When dislocations encounter a grain boundary, they pile up there. Slip can transmit to the adjacent grain only when the stress field ahead of the pileup is high enough. Direct observation of the process using transmission electron microscopy (TEM) also shows that the main mechanisms for the aforementioned slip transmission are the dislocation absorption and reemission for the low-angle boundaries (Soer et al. 2005) and the dislocation nucleation in the adjacent grain for the high-angle boundaries (Ohmura et al. 2004), respectively. As soon as deformation initiates in the adjacent grain, the grain boundary begins to deform and the plastic strain on the grain boundary increases. The energy associated with the deformation of the grain boundary in this case is taken to be mainly due to the energy dissipation as dislocations move in the grain boundary region. In addition to considering the resistance force to dislocation motion being temperature and rate dependent, this energy dissipation can be taken as a linear function of GB plastic strain.

Moreover, change in the grain boundary area can also affect the energy dissipation. The macroscopic accumulated plastic strain at the grain boundary, e^{pGB} , can be related to the microscopically deformation of the grain boundary through the root-mean-square of the gradient of this deformation. In addition, the energy change after the grain boundary has yielded, i.e., onset of slip transmission, can be approximated by a quadratic function of the aforementioned displacement gradient at microscale and hence the GB plastic strain at macroscale.

Combining both aforementioned mechanisms, i.e., change in the grain boundary area and deformation of the grain boundary due to the dislocation movement, involved in the energy dissipation due to the plastic strain transfer across the grain boundary, one can postulate the following generalized expression for the GB dissipation potential:

$$\mathcal{D}^{GB} = \frac{\ell_{dis}^{GB}}{m^{GB} + 1} \left(\sigma_0^{GB} + \mathcal{H}_0^{GB} e^{pGB(post)} \right) \left(1 - \frac{\mathcal{T}^{GB}}{\mathcal{T}_y^{GB}} \right)^{n^{GB}} \left(\frac{\dot{e}^{pGB(post)}}{\dot{\mathcal{p}}^{GB}} \right)^{m^{GB}} \dot{e}^{pGB(post)} \geq 0 \quad (80)$$

where ℓ_{dis}^{GB} is the GB dissipative length scale, m^{GB} and $\dot{\mathcal{p}}^{GB}$ are the viscous related material parameters, σ_0^{GB} is a constant accounting for the GB yield stress, \mathcal{H}_0^{GB} is the GB hardening parameter, \mathcal{T}_y^{GB} is the scale-independent GB thermal parameter at the onset of yield, and n^{GB} is the GB thermal parameter. The temperature and rate dependencies of the GB energy are shown, respectively, in the terms $\left(1 - \mathcal{T}^{GB} / \mathcal{T}_y^{GB} \right)^{n^{GB}}$ and $\left(\dot{e}^{pGB(post)} / \dot{\mathcal{p}}^{GB} \right)^{m^{GB}}$.

By using Eqs. (77) and (80), the GB dissipative microforce $\mathbb{M}^{GB,dis}$ can be obtained as:

$$\mathbb{M}^{GB,dis} = \ell_{dis}^{GB} \left(\sigma_0^{GB} + \mathcal{H}_0^{GB} e^{pGB(post)} \right) \left(1 - \frac{\mathcal{T}^{GB}}{\mathcal{T}_y^{GB}} \right)^{n^{GB}} \left(\frac{\dot{e}^{pGB(post)}}{\dot{\mathcal{p}}^{GB}} \right)^{m^{GB}} \quad (81)$$

Therefore, the GB thermodynamic microforce \mathbb{M}^{GB} can be obtained as:

$$\mathbb{M}^{GB} = G \ell_{en}^{GB} e^{pGB(pre)} + \ell_{dis}^{GB} \left(\sigma_0^{GB} + \mathcal{H}_0^{GB} e^{pGB(post)} \right) \left(1 - \frac{\mathcal{T}^{GB}}{\mathcal{T}_y^{GB}} \right)^{n^{GB}} \left(\frac{\dot{e}^{pGB(post)}}{\dot{\mathcal{p}}^{GB}} \right)^{m^{GB}} \quad (82)$$

Flow Rule: Grain Boundary

The flow rule for the grain boundary can be derived by substituting Eq. (82) into the microforce balances for the grain boundary, Eq. (70), such as: for S^{GBG_1} ,

$$\left[\sigma_0 \ell_{en}^2 \left[\ell_{en}^2 \left(e_{.k}^p e_{.k}^p \right) \right]^{\frac{\nu-1}{2}} e_{.i}^p + \sigma_0 \ell_{dis}^2 (m_2 + 1) \left[1 - \left(\frac{\mathcal{T}}{\mathcal{T}_y} \right)^n \right] \left(\frac{\dot{\mathcal{p}}}{\dot{\mathcal{p}}_2} \right)^{m_2} \frac{\dot{e}_{.i}^p}{\dot{\mathcal{p}}} \right] n_k^{GB} + G \ell_{en}^{GB} e^{pGB(pre)} = -\ell_{dis}^{GB} \left(\sigma_0^{GB} + \mathcal{H}_0^{GB} e^{pGB(post)} \right) \left(1 - \frac{\mathcal{T}^{GB}}{\mathcal{T}_y^{GB}} \right)^{n^{GB}} \left(\frac{\dot{e}^{pGB(post)}}{\dot{\mathcal{p}}^{GB}} \right)^{m^{GB}} \quad (83)$$

for S^{GBG_2} ,

$$\begin{aligned} & \left[\sigma_0 \ell_{en}^2 \left[\ell_{en}^2 \left(e_{,k}^p e_{,k}^p \right) \right]^{\frac{\theta-1}{2}} e_{,i}^p + \sigma_0 \ell_{dis}^2 (m_2 + 1) \left[1 - \left(\frac{\mathcal{J}}{\mathcal{J}_y} \right)^n \right] \left(\frac{\dot{\mathcal{P}}}{\dot{\mathcal{P}}_2} \right)^{m_2} \frac{\dot{e}_{,i}^p}{\dot{\mathcal{P}}} \right] n_k^{GB} \\ & - G \ell_{en}^{GB} e^{p^{GB(pre)}} = \ell_{dis}^{GB} \left(\sigma_0^{GB} + \mathcal{H}_0^{GB} e^{p^{GB(post)}} \right) \\ & \left(1 - \frac{\mathcal{J}^{GB}}{\mathcal{J}_y^{GB}} \right)^{n^{GB}} \left(\frac{\dot{e}^{p^{GB(post)}}}{\dot{\mathcal{P}}^{GB}} \right)^{m^{GB}} \end{aligned} \quad (84)$$

where the second term in LHS of both equations represents the backstress. Note that, in general case, the grain boundary model parameters are not identical on each side; however in this chapter, the same values are assumed to be considered for simplification.

Considering the GB flow rules as the boundary conditions of the grain interior flow rule, Eq. (62), results in a yield condition accounting for the temperature- and rate-dependent barrier effect of grain boundaries on the plastic slip and consequently the influence on the GND evolution in the grain interior.

Finite Element Formulation of the Proposed SGP Model

A two-dimensional finite element model for the proposed SGP model is developed to account for the size-dependent response for microscopic structures. The boundary value problem consists of solving the flow rules for the grain interior/boundary given in sections “[Flow Rule: Grain Interior](#)” and “[Flow Rule: Grain Boundary](#)” in conjunction with the constitutive equations given in sections “[Energetic and Dissipative Thermodynamic Microforces: Grain Interior](#)” and “[Energetic and Dissipative Thermodynamic Microforces: Grain Boundary](#)” subjected to the prescribed displacement conditions u_i^\dagger on part of the boundary $\partial\Omega'_0$ and traction free condition on the remaining boundary part of the body. The microscopic and macroscopic force balances can then be described in the global weak form after utilizing the principle of virtual power and applying the corresponding boundary conditions, i.e., arbitrary virtual displacement fields $\delta u = 0$ on $\partial\Omega'_0$ and arbitrary virtual plastic strain fields $\delta e^p = 0$ on $\partial\Omega''_0$ as follows:

$$\int_{\Omega_0} (\sigma_{ij} \delta u_{i,j}) dV = 0 \quad (85)$$

$$\int_{\Omega_0} [(x - \bar{\sigma}) \delta e^p + Q_i \delta e_{,i}^p] dV = 0 \quad (86)$$

where $\bar{\sigma}$ is the resolved shear stress defined by

$$\bar{\sigma} = \sigma_{ij} N_{ij} = \bar{\sigma}_{ij} N_{ij} \tag{87}$$

The UEL subroutine in the finite element software ABAQUS/Standard (2012) is built in this chapter for numerically solving the global weak forms of macroscopic and microscopic force balances, Eqs. (85) and (86), respectively. In this finite element formulation, the plastic strain field e^p and the displacement field u_i are discretized independently, and both of the fields are taken as fundamental unknown nodal degrees of freedom. In this regard, the displacement and corresponding strain field, ε_{ij} , and the plastic strain and corresponding plastic strain gradient field $e^p_{,j}$ are obtained by using the interpolation as follows:

$$u_i = \sum_{\eta=1}^{n_u} \mathcal{U}_{u_i}^\eta \mathbb{N}^\eta \tag{88}$$

$$\varepsilon_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) = \frac{1}{2} \sum_{\eta=1}^{n_u} \left(\mathcal{U}_{u_i}^\eta \frac{\partial \mathbb{N}^\eta}{\partial x_j} + \mathcal{U}_{u_j}^\eta \frac{\partial \mathbb{N}^\eta}{\partial x_i} \right) \tag{89}$$

$$e^p = \sum_{\eta=1}^{n_{e^p}} \mathcal{E}_{e^p}^\eta \mathbb{N}^\eta \tag{90}$$

$$e^p_{,j} = \frac{\partial e^p}{\partial x_j} = \sum_{\eta=1}^{n_{e^p}} \mathcal{E}_{e^p}^\eta \frac{\partial \mathbb{N}^\eta}{\partial x_j} \tag{91}$$

where \mathbb{N}_u^η and $\mathbb{N}_{e^p}^\eta$ denote the interpolation functions and $\mathcal{U}_{u_i}^\eta$ and $\mathcal{E}_{e^p}^\eta$ denote the nodal values of the plastic strains and displacements at node η , respectively. The terms n_u and n_{e^p} represent the number of nodes per single element for displacement and plastic strain, respectively. Since a two-dimensional quadratic 9-node element is used in this chapter, n_u and n_{e^p} are set up as nine. It should be noted that n_u and n_{e^p} do not necessarily have to be the same as each other in the present finite element implementation, even though both the displacement and plastic strain fields are calculated by using the standard isoparametric interpolation functions.

The body is approximated using finite elements, $\Omega = \cup \Omega_{el}$. By substituting Eqs. (88), (89), (90), and (91) into Eqs. (85) and (86), the nodal residuals for the displacement \mathbb{R}_{u_i} and the plastic strain \mathbb{R}_{e^p} for each element Ω_{el} can be obtained as:

$$(\mathbb{R}_{u_i})_\eta = - \int_{\Omega_{el}} \left(\sigma_{ij} \frac{\partial \mathbb{N}_u^\eta}{\partial x_j} \right) dV \tag{92}$$

$$(\mathbb{R}_{e^p})_\eta = - \int_{\Omega_{el}} \left[(x - \bar{\sigma}) \mathbb{N}_{e^p}^\eta + Q_i \frac{\partial \mathbb{N}_{e^p}^\eta}{\partial x_i} \right] dV \tag{93}$$

The system of linear equations, $(\mathbb{R}_{u_i})_\eta = 0$ and $(\mathbb{R}_{e^p})_\eta = 0$, are solved using ABAQUS/Standard (2012) based on the Newton-Raphson iterative method. Occasionally, the modified Newton-Raphson method, referred to as quasi-Newton-Raphson method, is employed in the case that the numerical solution suffers a divergence during the initial increment immediately after an abrupt change in loading. In the quasi-Newton-Raphson method, a specific correction factor, which is less than one, is multiplied by one portion of the stiffness matrix. By using this method, a divergence problem can be overcome; however, convergence is expected to be slow because of the expensive computational cost. The Taylor expansion of the residuals with regard to the current nodal values can be expressed by assuming the nodal displacement and the plastic strain in iteration ζ as $\mathcal{U}_{u_i}^\zeta$ and $\mathcal{E}_{e^p}^\zeta$ are, respectively, as follows:

$$\begin{aligned} (\mathbb{R}_{u_i} |_{\mathcal{U}_{u_i}^{\zeta+1}, \mathcal{E}_{e^p}^{\zeta+1}})_\eta &= (\mathbb{R}_{u_i} |_{\mathcal{U}_{u_i}^\zeta, \mathcal{E}_{e^p}^\zeta})_\eta + \left(\frac{\partial \mathbb{R}_{u_i}}{\partial \mathcal{U}_{u_k}^\eta} \Big|_{\mathcal{U}_{u_i}^\zeta} \right) \Delta \mathcal{U}_{u_i}^\eta \\ &+ \left(\frac{\partial \mathbb{R}_{u_i}}{\partial \mathcal{E}_{e^p}^\eta} \Big|_{\mathcal{E}_{e^p}^\zeta} \right) \Delta \mathcal{E}_{e^p}^\eta + \mathcal{O} \left((\Delta \mathcal{U}_{u_i}^\eta)^2, (\Delta \mathcal{E}_{e^p}^\eta)^2 \right) \end{aligned} \quad (94)$$

$$\begin{aligned} (\mathbb{R}_{e^p} |_{\mathcal{U}_{u_i}^{\zeta+1}, \mathcal{E}_{e^p}^{\zeta+1}})_\eta &= (\mathbb{R}_{e^p} |_{\mathcal{U}_{u_i}^\zeta, \mathcal{E}_{e^p}^\zeta})_\eta + \left(\frac{\partial \mathbb{R}_{e^p}}{\partial \mathcal{U}_{u_k}^\eta} \Big|_{\mathcal{U}_{u_i}^\zeta} \right) \Delta \mathcal{U}_{u_i}^\eta \\ &+ \left(\frac{\partial \mathbb{R}_{e^p}}{\partial \mathcal{E}_{e^p}^\eta} \Big|_{\mathcal{E}_{e^p}^\zeta} \right) \Delta \mathcal{E}_{e^p}^\eta + \mathcal{O} \left((\Delta \mathcal{U}_{u_i}^\eta)^2, (\Delta \mathcal{E}_{e^p}^\eta)^2 \right) \end{aligned} \quad (95)$$

where $\Delta \mathcal{U}_{u_i}^\eta = (\mathcal{U}_{u_i}^{\zeta+1})_\eta - (\mathcal{U}_{u_i}^\zeta)_\eta$, $\Delta \mathcal{E}_{e^p}^\eta = (\mathcal{E}_{e^p}^{\zeta+1})_\eta - (\mathcal{E}_{e^p}^\zeta)_\eta$, and $\mathcal{O} \left((\Delta \mathcal{U}_{u_i}^\eta)^2, (\Delta \mathcal{E}_{e^p}^\eta)^2 \right)$ is the big O notation to represent the terms of higher order than the second degree. These residuals are repetitively calculated at every time step, and the calculated numerical results are updated during the whole iterations. The increments in the nodal displacement and plastic strains can be obtained by computing the system of linear equations shown in Eq. (96):

$$\underbrace{\begin{bmatrix} K_{u_i u_k}^{\Omega_{el}} & K_{u_i e^p}^{\Omega_{el}} \\ K_{e^p u_k}^{\Omega_{el}} & K_{e^p e^p}^{\Omega_{el}} \end{bmatrix}}_{\mathbf{K}^{\Omega_{el}}} \begin{Bmatrix} \Delta \mathcal{U}_{u_k}^\eta \\ \Delta \mathcal{E}_{e^p}^\eta \end{Bmatrix} = \begin{Bmatrix} (\mathbb{R}_{u_i} |_{\mathcal{U}_{u_i}^\zeta, \mathcal{E}_{e^p}^\zeta})_\eta \\ (\mathbb{R}_{e^p} |_{\mathcal{U}_{u_i}^\zeta, \mathcal{E}_{e^p}^\zeta})_\eta \end{Bmatrix} \quad (96)$$

where $\mathbf{K}^{\Omega_{el}}$ is the Jacobian (stiffness) matrix.

From the functional forms of the thermodynamic microforces defined in sections “Energetic and Dissipative Thermodynamic Microforces: Grain Interior” and “Energetic and Dissipative Thermodynamic Microforces: Grain Boundary” and Eqs. (94) and (95) along with the Eqs. (88), (89), (90), and (91) at the end of each time step, each component of the Jacobian matrix can be obtained, respectively, as

follows:

$$K_{u_i u_k}^{\Omega_{el}} = - \frac{\partial \mathbb{R}_{u_i}}{\partial \mathcal{U}_{u_k}^\eta} \Big|_{\mathcal{U}_{u_i}^\zeta} = \int_{\Omega_{el}} \left(E_{ijkl} \frac{\partial \mathbb{N}_u^\eta}{\partial x_j} \frac{\partial \mathbb{N}_u^\eta}{\partial x_l} \right) dV \quad (97)$$

$$K_{u_i e^p}^{\Omega_{el}} = - \frac{\partial \mathbb{R}_{u_i}}{\partial \mathcal{E}_{e^p}^\eta} \Big|_{\mathcal{E}_{e^p}^\zeta} = \int_{\Omega_{el}} \left(E_{ijkl} \frac{e^p}{\varepsilon_{kl}} \frac{\partial \mathbb{N}_u^\eta}{\partial x_j} \mathbb{N}_{e^p}^\eta \right) dV \quad (98)$$

$$K_{e^p u_k}^{\Omega_{el}} = - \frac{\partial \mathbb{R}_{e^p}}{\partial \mathcal{U}_{u_k}^\eta} \Big|_{\mathcal{U}_{u_i}^\zeta} = \int_{\Omega_{el}} \left(E_{ijkl} \frac{e^p}{\varepsilon_{ij}} \mathbb{N}_u^\eta \frac{\partial \mathbb{N}_{e^p}^\eta}{\partial x_l} \right) dV \quad (99)$$

$$\begin{aligned} K_{e^p e^p}^{\Omega_{el}} &= - \frac{\partial \mathbb{R}_{e^p}}{\partial \mathcal{E}_{e^p}^\eta} \Big|_{\mathcal{E}_{e^p}^\zeta} \\ &= \int_{\Omega_{el}} \left[\left(r \mathcal{H}_0(e^p)^r + \mathcal{H}_0 \left(\frac{\dot{e}^p}{\dot{\rho}_1} \right)^{m_1} + \sigma_0 \sqrt{\mathcal{H}^2(e^p) + \ell_{N-G\mathcal{E}^p}} \frac{m_1 \dot{e}^p{}^{m_1-1}}{\dot{\rho}_1^{m_1} \Delta t} \right) \right. \\ &\quad \left[1 - \left(\frac{\mathcal{J}}{\mathcal{J}_y} \right)^n \right] \mathbb{N}_{e^p}^\eta \mathbb{N}_{e^p}^\eta + \sigma_0 \ell_{en}^2 \left[\ell_{en}^2 (e^p_k e^p_k) \right]^{\frac{\theta-1}{2}} \frac{\partial \mathbb{N}_{e^p}^\eta}{\partial x_j} \frac{\partial \mathbb{N}_{e^p}^\eta}{\partial x_j} + \sigma_0 \ell_{dis}^4 (m_2^2 - 1) \\ &\quad \left[1 - \left(\frac{\mathcal{J}}{\mathcal{J}_y} \right)^n \right] \frac{\dot{\rho}^{m_2-3}}{\dot{\rho}_2^{m_2} \Delta t} \left(\dot{e}^p_j \frac{\partial \mathbb{N}_{e^p}^\eta}{\partial x_j} \right) \left(\dot{e}^p_j \frac{\partial \mathbb{N}_{e^p}^\eta}{\partial x_j} \right) \\ &\quad + \sigma_0 \ell_{dis}^2 (m_2 + 1) \left[1 - \left(\frac{\mathcal{J}}{\mathcal{J}_y} \right)^n \right] \frac{\dot{\rho}^{m_2-1}}{\dot{\rho}_2^{m_2} \Delta t} \frac{\partial \mathbb{N}_{e^p}^\eta}{\partial x_j} \frac{\partial \mathbb{N}_{e^p}^\eta}{\partial x_j} \Big] dV \\ &\quad - \left[G \ell_{en}^{GB} + \ell_{dis}^{GB} \frac{(\sigma_0^{GB} + \mathcal{H}_0^{GB} e^p{}^{GB(post)})}{(\Delta t \dot{\rho}^{GB})^{m^{GB}}} \right. \\ &\quad \left. \left(1 - \frac{\mathcal{J}^{GB}}{\mathcal{J}_y^{GB}} \right)^{n^{GB}} \dot{e}^p{}^{GB(post)m^{GB}-1} \right] \mathbb{N}_{e^p}^\eta \mathbb{N}_{e^p}^\eta \end{aligned} \quad (100)$$

where Δt is a time step. The grain boundary terms in Eq. (100) are only applied for nodes on the grain boundary area.

Validation of the Proposed SGP Model

In this section, the proposed SGP theory and corresponding finite element implementation are validated through the comparison against the experimental measurements from two sets of size effect tests: aluminum thin film experiments by Haque and Saif (2003) and nickel thin film experiments by Han et al. (2008). The material parameters for two different metals, aluminum and nickel, are also calibrated by

using the experimental data. The parameters σ_0 and \mathcal{H}_0 are determined by extrapolating the experimental data, and the material length scales are determined based on the suggestion by Anand et al. (2005). In their work, an initial assumption for the material length scales is suggested by matching the numerical results from the proposed flow rule to the yield strength and backstress experimental measurements under the assumption of $\ell_{en} = 0$ and $\ell_{dis} = 0$ at each case, respectively. The rest of the material parameters come from the literature (Han et al. 2008; Haque and Saif 2003; Voyiadjis and Song 2017). The grain boundary flow rule is not considered in sections “Validation of the Proposed SGP Model,” “Microfree Grain Boundary,” and “Microhard Grain Boundary.”

Haque and Saif (2003) developed the micro-electromechanical system (MEMS)-based testing skill for the nanoscale aluminum (Al) thin films under the uniaxial tensile loading to investigate the strain gradient effect in 100 nm, 150 nm, 200 nm, and 485 nm-thick specimens, which have the average grain sizes of 50 nm, 65 nm, 80 nm, and 212 nm, respectively. The specimens with 99.99% pure sputter-deposited freestanding Al thin films are 10 μm wide and 275 μm long. All experiments are carried out in situ in SEM, and the strain and stress resolutions for the tests are set as 0.03% and 5 MPa, respectively. The general and calibrated material parameters are shown in Table 1. Figure 1 displays the direct comparison between the proposed model and the experimental observations. As clearly shown in this figure, the size effect, “smaller is stronger,” is observed on the stress-strain

Table 1 The general and calibrated material parameters used for the validation of the proposed strain gradient plasticity model. (Reprinted with permission from Song and Voyiadjis (2018a))

General		Aluminum	Nickel
E (GPa)	Elastic modulus for isotropic linear elasticity	70	115
ν	Poisson’s ratio	0.30	0.31
μ (GPa)	Shear modulus for isotropic linear elasticity	27	44
ρ ($\text{g} \cdot \text{cm}^{-3}$)	Density	2.702	8.902
c_ε ($\text{J/g} \cdot ^\circ\text{K}$)	Specific heat capacity at constant stress	0.910	0.540
α^{th} ($\mu\text{m/m} \cdot ^\circ\text{K}$)	Thermal expansion coefficient	24.0	13.1
$\dot{\phi}_1, \dot{\phi}_2$ (s^{-1})	Reference plastic strain rate	0.04	0.04
τ	Nonlinear hardening material constant	0.6	0.2
m_1	Nonnegative strain rate sensitivity parameter	0.05	0.05
m_2	Nonnegative strain rate sensitivity parameter	0.2	0.2
T_y ($^\circ\text{K}$)	Thermal material parameter	933	890
n	Temperature sensitivity parameter	0.3	0.3
Calibrated		Aluminum	Nickel
σ_0 (MPa)	Stress-dimensioned scaling constant	1,000	950
\mathcal{H}_0 (MPa)	Isotropic hardening parameter	1,000	1,000
ℓ_{en} (μm)	Energetic length scale	1.2	1.6
ℓ_{dis} (μm)	Dissipative length scale	0.7	0.5
ℓ_{N-G} (μm)	N-G length scale	1.0	2.0

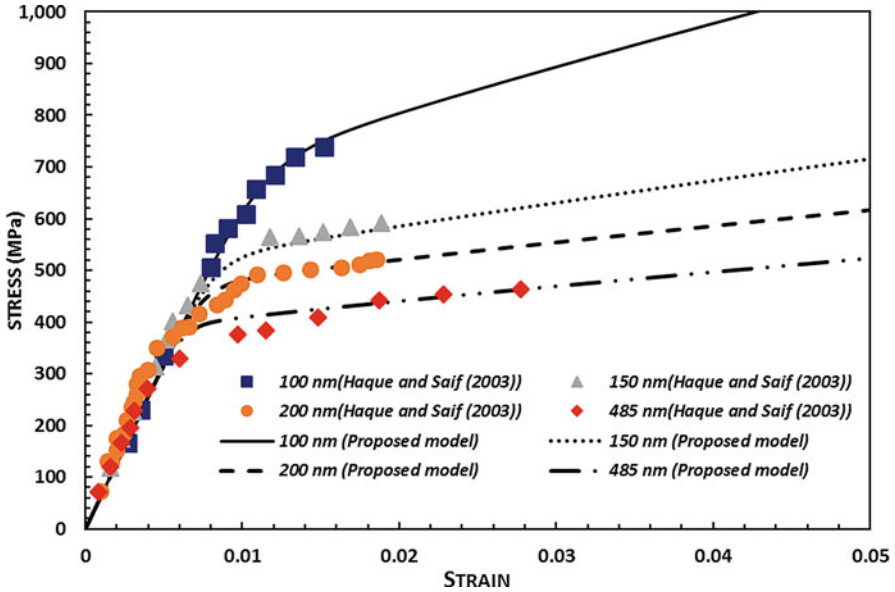
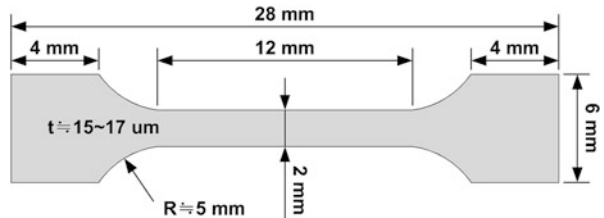


Fig. 1 The validation of the proposed SGP model by comparing the numerical results from the proposed model with the experimental measurements from Haque and Saif (2003) on the stress-strain response of the sputter-deposited Al thin films. (Reprinted with permission from Song and Voyiadjis (2018a))

Fig. 2 The schematic illustration of the dog bone specimen and its main dimensions (Han et al. 2008). (Reprinted with permission from Song and Voyiadjis (2018a))



curves; furthermore, the numerical results from the proposed SGP model and the experimental data correspond closely with each other.

The microscale tensile experiment skill for evaluating the mechanical and thermal properties of the nickel (Ni) thin films at high temperatures is developed by Han et al. (2008). The dog bone-shaped specimens used in their experiments were made by micro-electromechanical system (MEMS) processes, and the primary dimensions of the specimen are given in Fig. 2. The general and calibrated material parameters for Ni are presented in Table 1. The experimental measurements at four different temperatures, i.e., 25°C, 75°C, 145°C, and 218°C, and corresponding numerical values from the proposed model are presented in Fig. 3. As shown in this figure, it is obvious from both the numerical and experimental results that the Young’s modulus is not affected by variations in temperature

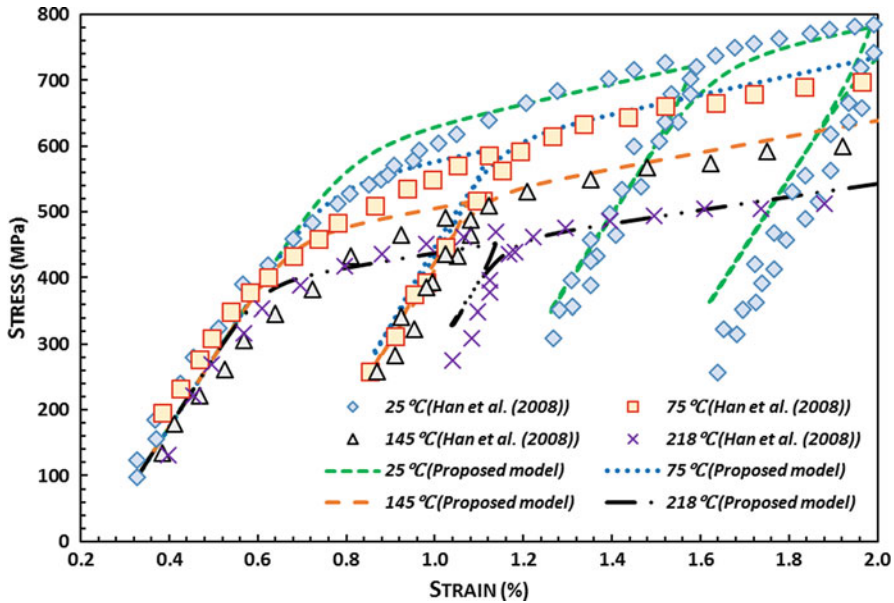
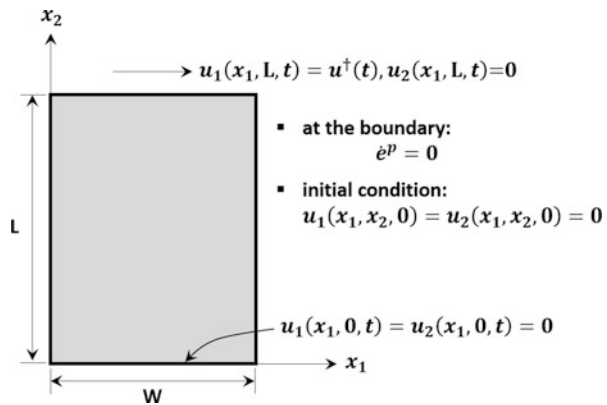


Fig. 3 The validation of the proposedSGP model by comparing the numerical results from the proposed model with the experimental measurements from Han et al. (2008) on the stress-strain response of the Ni thin films. (Reprinted with permission from Song and Voyiadjis (2018a))

Fig. 4 The schematic illustration of the simple shear problem including the macroscopic and microscopic boundary conditions and initial conditions. (Reprinted with permission from Song and Voyiadjis (2018a))



while the yield strength decreases as the specimen temperature increases; in addition, Fig. 3 clearly shows that the Bauschinger effect is not affected very much by variations in the specimen temperature. All these observations are not much different from those observed in one-dimensional finite element simulations (Voyiadjis and Song 2017).

Microfree Grain Boundary

The characteristics of the proposed SGP theory under the microfree boundary condition at the grain boundary is addressed in this section by solving the shear problem of a rectangular plate with varying material parameters.

By following Gurtin (2003), the simple class of microscopically free boundary condition on a prescribed subsurface is employed from Eq. (14) as follows:

$$m = Q_i n_i = 0 \quad (101)$$

The schematic illustration of the problem, the initial conditions, and the macroscopic and microscopic boundary conditions are shown in Fig. 4. The parameter $u^\dagger(t)$ represents the prescribed displacement. The stress-strain behavior and the distributions of the temperature and accumulated plastic strain across the plate in x_2 direction are investigated for the various material length scales (ℓ_{en} , ℓ_{dis} , and ℓ_{N-G}), the hardening parameter \mathcal{H}_0 , and the temperature-related parameters (n and \mathcal{T}_y). The material parameters in Table 2 are used for this section unless stated otherwise.

Energetic Gradient Hardening

For exploring the characteristics of the energetic gradient hardening only, the dissipative gradient and GND gradient terms are assumed to vanish by imposing $\ell_{dis} = \ell_{N-G} = 0$. The stress-strain behaviors and the distributions of e^p and \mathcal{T}

Table 2 The material parameters used in sections “Microfree Grain Boundary” and “Microhard Grain Boundary.” (Reprinted with permission from Song and Voyiadjis (2018a))

General		Values
E (GPa)	Elastic modulus for isotropic linear elasticity	110
ν	Poisson's ratio	0.33
μ (GPa)	Shear modulus for isotropic linear elasticity	48
ρ ($g \cdot cm^{-3}$)	Density	8.960
c_ε ($J/g \cdot ^\circ K$)	Specific heat capacity at constant stress	0.385
α^{th} ($\mu m/m \cdot ^\circ K$)	Thermal expansion coefficient	24.0
$\dot{\rho}_1, \dot{\rho}_2$ (s^{-1})	Reference plastic strain rate	0.04
\mathcal{r}	Nonlinear hardening material constant	0.6
m_1	Nonnegative strain rate sensitivity parameter	0.05
m_2	Nonnegative strain rate sensitivity parameter	0.2
\mathcal{T}_y ($^\circ K$)	Thermal material parameter	1,358
n	Temperature sensitivity parameter	0.3
σ_0 (MPa)	Stress-dimensioned scaling constant	1,000
\mathcal{H}_0 (MPa)	Isotropic hardening parameter	1,000

along the height of the plate for $\ell_{en}/L = 0, 0.1, 0.2, 0.3, 0.5, 0.7,$ and 1.0 with $\mathcal{H}_0 = 0$ MPa are shown in Fig. 5. For the case of $\ell_{en}/L = 0$, the stress-strain curve does not show hardening as expected, and the distribution of e^p is uniform. As ℓ_{en}/L increases, the rates of the hardening increase as shown in Fig. 5a, and the distributions of e^p and \mathcal{T} become more parabolic. When the dissipative length scale ℓ_{dis} is absent, the heat generation is governed by the amount of e^p ; therefore the distribution of \mathcal{T} is identical to the one of e^p as observed in Fig. 5b, c. Due to the microhard boundary condition, the temperature does not rise at the boundary, and the maximum temperature is observed at the center of the plate.

Dissipative Gradient Strengthening

It is now assumed that the energetic and GND hardening disappear by setting $\ell_{en} = \ell_{N-G} = 0$. The isotropic hardening parameter is also set as $\mathcal{H}_0 = 0$ MPa. The stress-strain curves and the distributions of e^p and \mathcal{T} across the height of the plate for $\ell_{dis}/L = 0, 0.05, 0.1, 0.2, 0.3, 0.5,$ and 0.7 are shown in Fig. 6. As ℓ_{dis}/L increases, it is shown in Fig. 6a that the initial yield strength increases without strain hardening. The non-monotonic behavior of e^p with ℓ_{dis}/L is elucidated by plotting the maximal values of e^p at the center of the plate for varying ℓ_{dis}/L as shown in Fig. 6c. These numerical results are in line with the works of Bardella (2006) and Fleck and Hutchinson (2001). The drastic change shown in the accumulated plastic strain distribution according to varying ℓ_{dis}/L prominently affects the temperature distribution. According to Eq. (67), both the plastic strain and its gradient are involved in the evolution of \mathcal{T} . As observed from Fig. 6b, the large amounts of plastic strain gradient raise the temperature at the boundary; on the other hand, the temperature at the center of the plate is primarily influenced by plastic strain.

GND Hardening

It is now assumed that the energetic and dissipative gradient terms disappear by setting $\ell_{en} = \ell_{dis} = 0$. The stress-strain curves and the distributions of e^p and \mathcal{T} across the height of the plate for $\ell_{N-G}/L = 0, 0.1, 0.2, 0.5, 1.0, 1.5,$ and 2.0 with $\mathcal{H}_0 = 0$ MPa are shown in Fig. 7. As ℓ_{N-G}/L increases, the rates of the strain hardening increase more significantly, while no initial yield strength is observed. Similar to the energetic gradient hardening case, in the absence of ℓ_{dis} , the distribution of temperature is identical to the one of plastic strain.

Temperature-Related Parametric Study

The effects are now investigated of the two material parameters, n and \mathcal{T}_y , on the stress-strain behavior, the distributions of the temperature and accumulated plastic

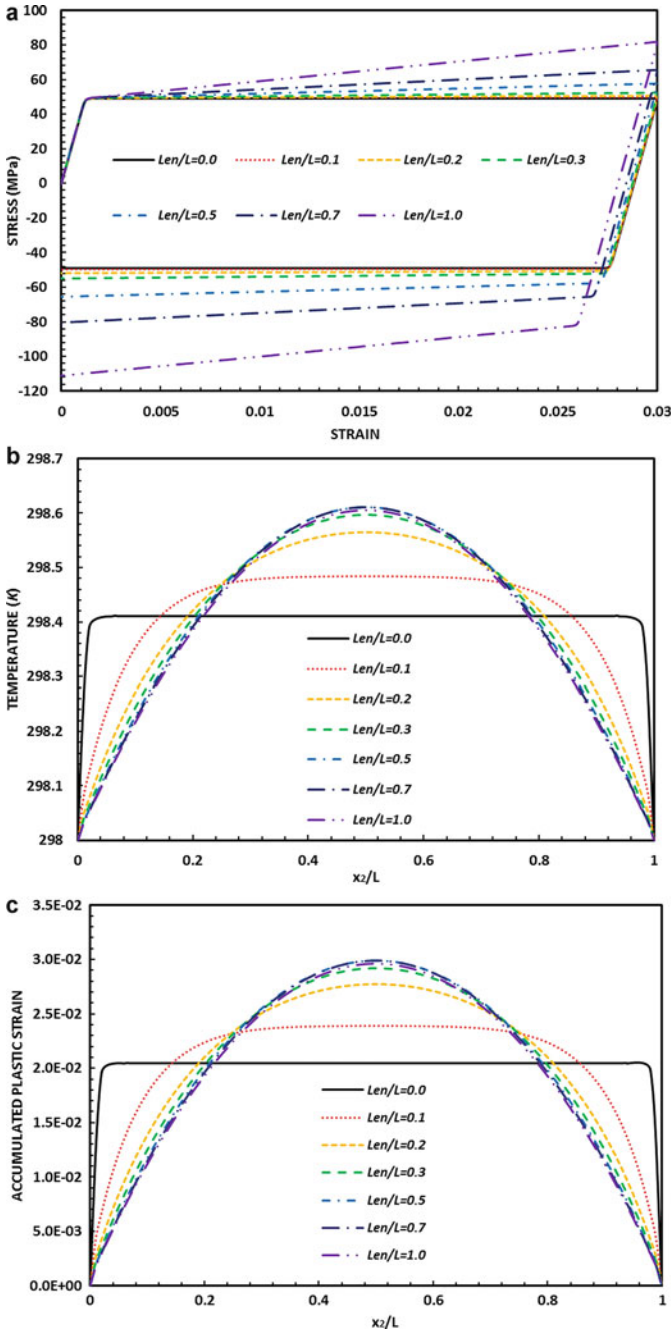


Fig. 5 The effects of the energetic gradient hardening with the energetic length scale only ($\ell_{dis} = \ell_N - G = 0$): (a) the stress-strain response, (b) the temperature distribution, and (c) the accumulated plastic strain distribution. (Reprinted with permission from Song and Voyiadjis (2018a))

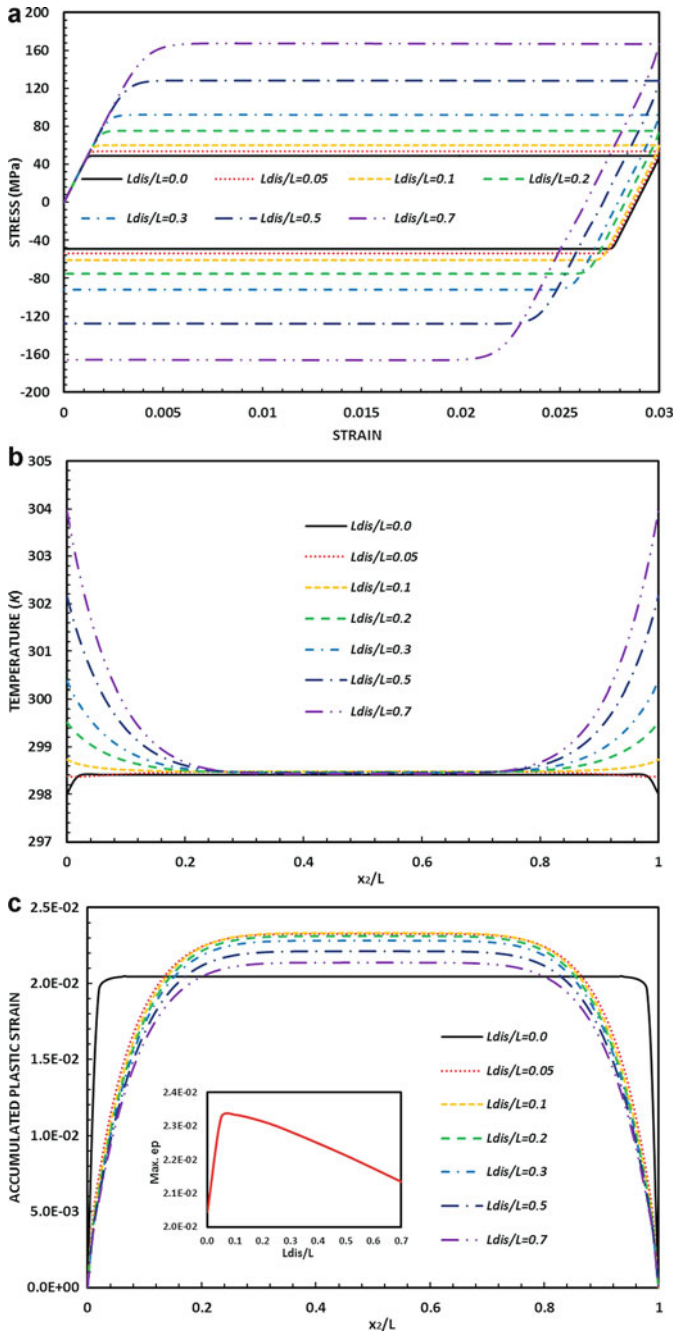


Fig. 6 The effects of the dissipative gradient strengthening with the dissipative length scale only ($\ell_{en} = \ell_{N-G} = 0$): (a) the stress-strain response, (b) the temperature distribution, and (c) the accumulated plastic strain distribution. (Reprinted with permission from Song and Voyiadjis (2018a))

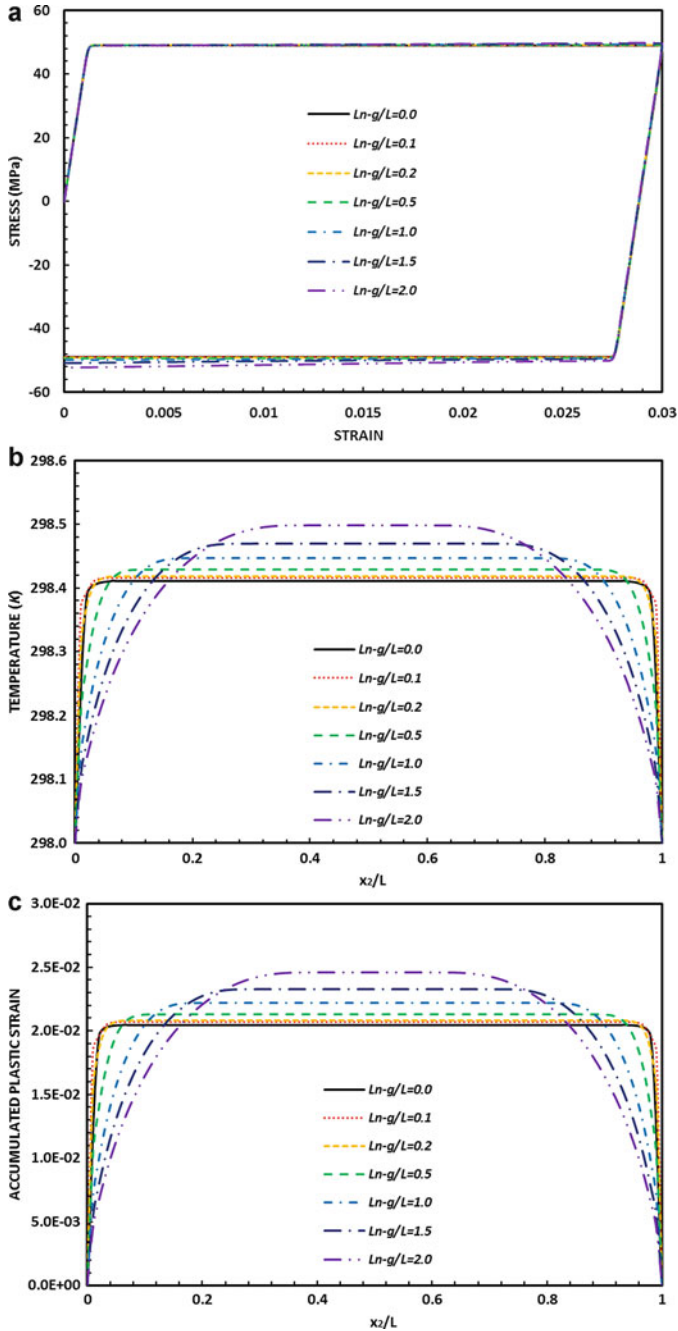


Fig. 7 The effects of the GND hardening with the N-G length scale only ($\ell_{en} = \ell_{dis} = 0$): (a) the stress-strain response, (b) the temperature distribution, and (c) the accumulated plastic strain distribution. (Reprinted with permission from Song and Voyiadjis (2018a))

strain, and the temperature evolution in conjunction with $\ell_{en}/L = 1.0$, $\ell_{dis}/L = 0.5$, and $\ell_{N-G}/L = 1.0$.

Figures 8 and 9 show the effect of the material parameters n and \mathcal{T}_y applying the temperature-dependent behavior of the present model on the stress-strain curve, the distributions of the temperature and accumulated plastic strain distribution, and the evolution of the temperature at the midpoint of the plate. As shown in these figures, the effects of the two parameters on the thermal and mechanical material response are similar to each other due to the fact that two parameters affect the temperature-related term $(1 - (\mathcal{T}/\mathcal{T}_y)^n)$ in the flow rule, which explains the thermal activation mechanism for overcoming the local barriers to the dislocation movement, in a similar way. It is worth mentioning that, in some literatures, the parameter \mathcal{T}_y is assumed to be equal to the melting temperature of the specific material in order not to bring in the material parameter additionally. In this chapter, however, the normal parameter is employed as a normalizing constant that should be calibrated by using the experimental data.

It is clearly observed in Figs. 8a and 9a that, as n and \mathcal{T}_y increase, the strain hardening and the initial yield strength also increase. The effect, nevertheless, is more prominent in initial yield strength, because strain hardening mechanism is affected by temperature through the dislocation forest barriers; on the other hand, the backstress does not depend on the temperature. For both n and \mathcal{T}_y , it is shown in Figs. 8b, c and 9b, c that the accumulated plastic strains at the boundary ($0 < x_2/L < 0.2$ and $0.8 < x_2/L < 1$) are almost identical with respect to the different values of n and \mathcal{T}_y ; on the other hand, the most significant differences in the temperature profiles occur at the same range for both parameters. In contrast to this, the most substantial differences in the accumulated plastic strain profiles according to the values of n and \mathcal{T}_y occur in the middle ($0.2 < x_2/L < 0.8$) of the plate; on the other hand, the temperature profiles at this part do not show considerable difference.

Microhard Grain Boundary

In this section, a microhard boundary condition, which describes that the dislocation movements are blocked completely at the grain boundary, is assumed to hold on the grain boundary as follows:

$$e^p = 0 \quad (102)$$

The square plate with an edge of W is solved to investigate the effect of this grain boundary condition. The top edge of the plate is subject to a prescribed condition in terms of the displacement ($u_1^{top}(x_1, W, t) = u_1^*(t)$, $u_2^{top}(x_1, W, t) = 0$), while the bottom edge is fixed ($u_1^{bot}(x_1, 0, t) = u_2^{bot}(x_1, 0, t) = 0$). The whole square is meshed using 1,600 (40×40) elements and split into several grains by the four different grain boundary areas, which are indicated by the bold lines as shown in Fig. 10. For these simulations, the stress-dimensioned scaling constant σ_0 and the

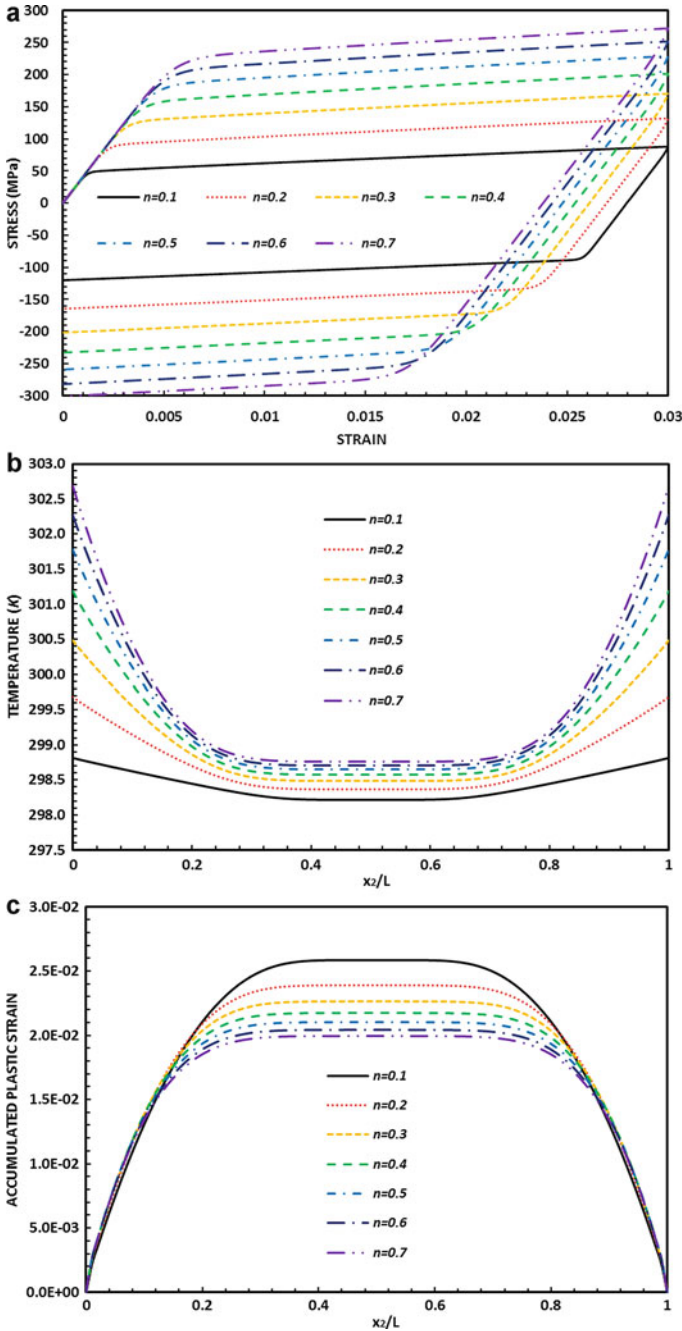


Fig. 8 The effects of the thermal material parameter n with $\ell_{en}/L = 1.0$, $\ell_{dis}/L = 0.5$, and $\ell_{N-G}/L = 1.0$: (a) the stress-strain response, (b) the temperature distribution, and (c) the accumulated plastic strain distribution. (Reprinted with permission from Song and Voyiadjis (2018a))

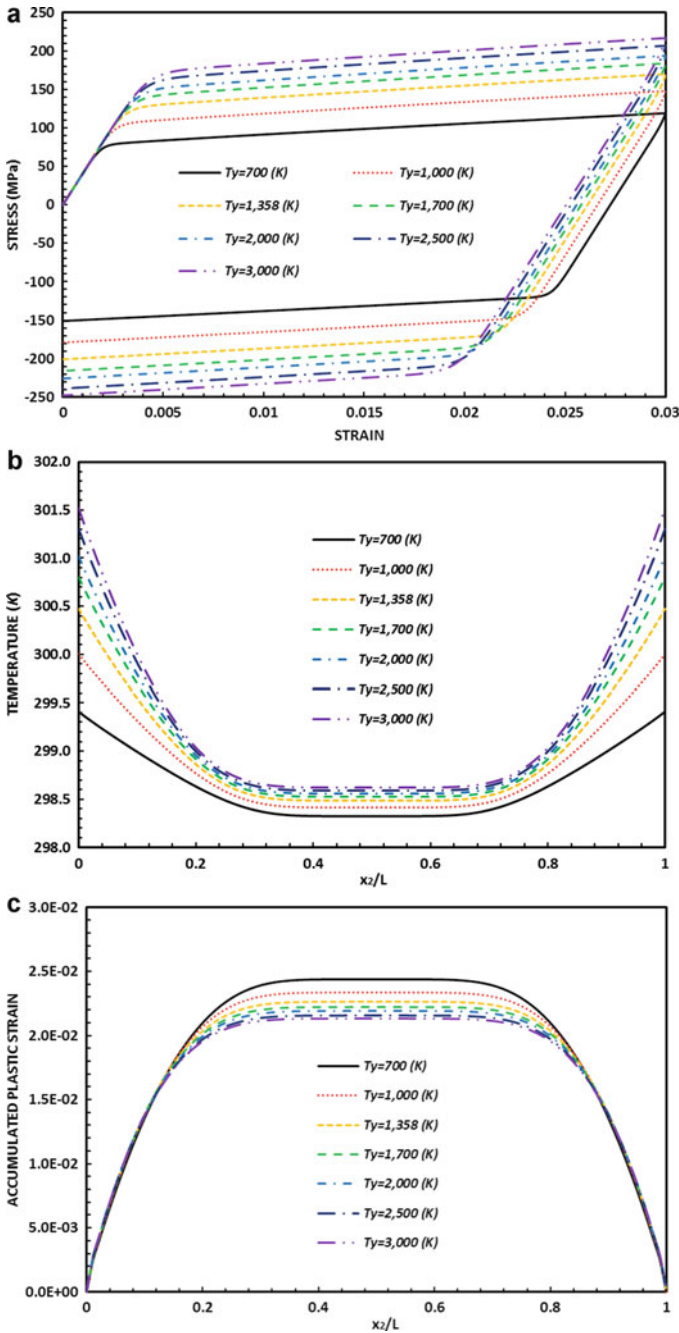


Fig. 9 The effects of the thermal material parameter T_y with $\ell_{em}/L = 1.0$, $\ell_{dis}/L = 0.5$, and $\ell_{N-G}/L = 1.0$: (a) the stress-strain response, (b) the temperature distribution, and (c) the accumulated plastic strain distribution. (Reprinted with permission from Song and Voyiadjis (2018a))

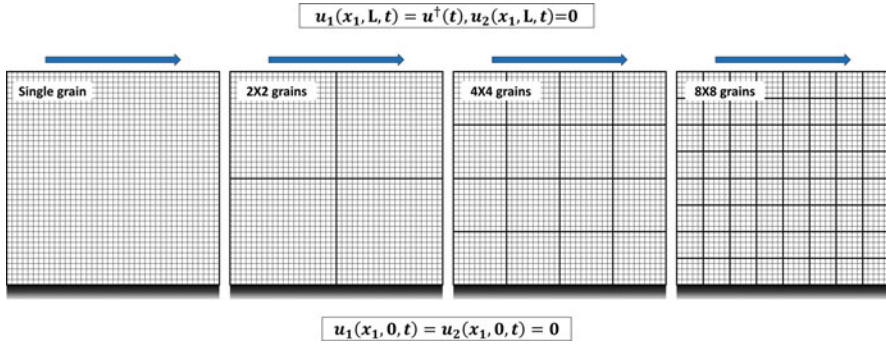


Fig. 10 The schematic illustration of the square plate problem with four different grain boundary areas. (Reprinted with permission from Song and Voyiadjis (2018a))

isotropic hardening parameter \mathcal{H}_0 are set as 195 MPa and 0 MPa, respectively, and the values in Table 2 are used again for the rest of the material parameters.

The comparison of the microscopic boundary condition is addressed through the single grain in Fig. 11. Figure 11a shows the distributions of the accumulated plastic strain e^p and temperature \mathcal{T} , respectively, with the microhard boundary condition at the grain boundaries, while Fig. 11b shows those with the microfree boundary condition. The terminologies “NT11” and “UVARM6” in Fig. 11 indicate the accumulated plastic strain and the temperature, respectively, and are used continually in the rest of this work. For this example, $\ell_{en}/L = 0.3$, $\ell_{dis}/L = 0.0$, and $\ell_{N-G}/L = 0.0$ are used. As can be seen in Fig. 11a, each edge of the plate with microhard boundary condition obstructs the dislocation movement, which results in $e^p = 0$, whereas in the case of microfree boundary condition, e^p and \mathcal{T} are evenly spread across the grain.

Figure 12 shows the distributions of e^p and \mathcal{T} with no gradient effect. This can be regarded as the reference case for the purpose of the comparison to other cases with the gradient effect.

The effects of the grain boundary area in conjunction with the energetic gradient hardening, dissipative gradient strengthening, GND hardening, and no gradient effect are shown in Fig. 13 through the stress-strain responses. As can be seen in Fig. 13a, no significant hardening or strengthening is observed under the conventional plasticity theory. On the other hand, by considering the gradient effect, the energetic hardening, dissipative strengthening, and GND hardening are observed, respectively, in Fig. 13b, c, d with varying grain boundary areas.

The effects of the grain boundary area in conjunction with the energetic gradient hardening, dissipative gradient strengthening, and GND hardening on the distributions of e^p and \mathcal{T} are also shown, respectively, in Figs. 14, 15, and 16. For the simulations in Fig. 14, the energetic length scales vary from 0.1 to 0.3, and the dissipative and N-G length scales are set as zero. It should be noted that the distributions of e^p and \mathcal{T} are not identical in all grains since the simple shear loading is applied to the top edge of the plate and not to each grain individually. It

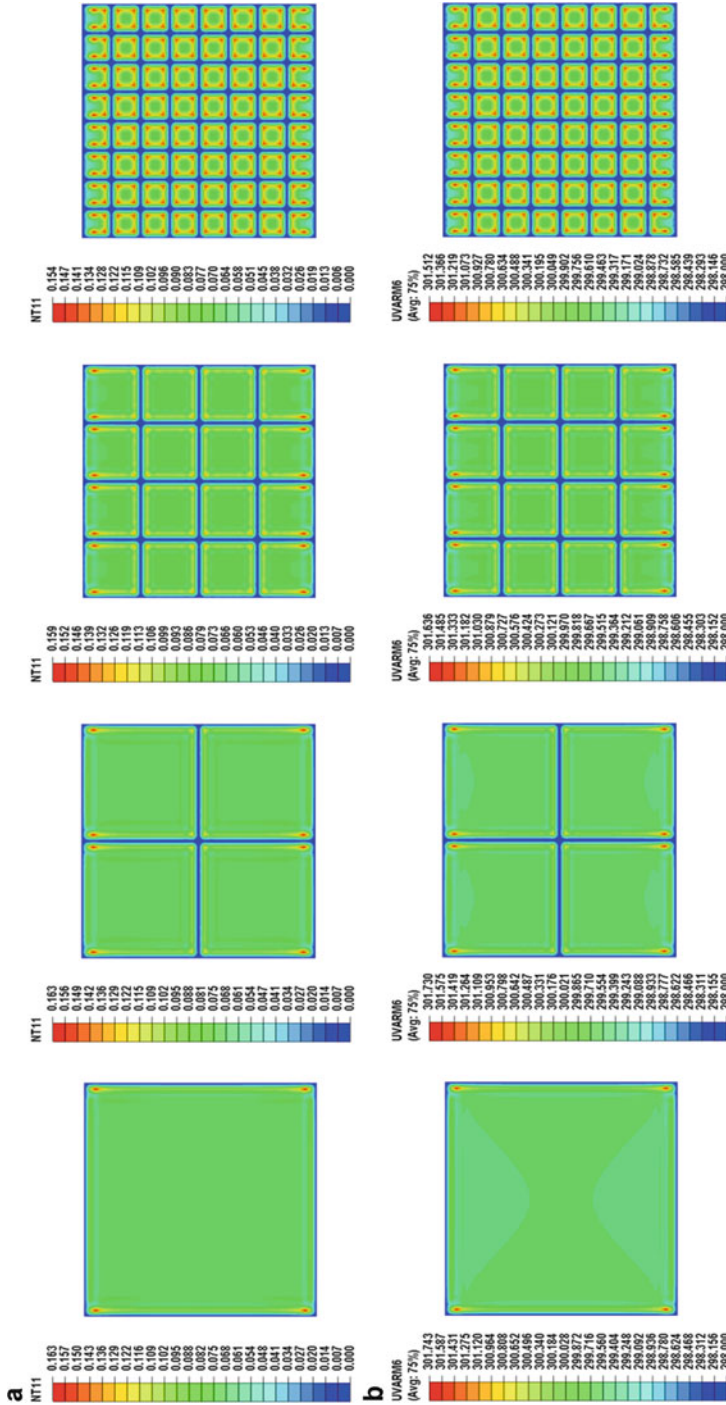


Fig. 12 The distributions of the accumulated plastic strain and the temperature in conjunction with varying grain boundary areas with $\ell_{em} = \ell_{dis} = \ell_{N-G} = 0$: (a) e^p and (b) T . (Reprinted with permission from Song and Voyiadjis (2018a))

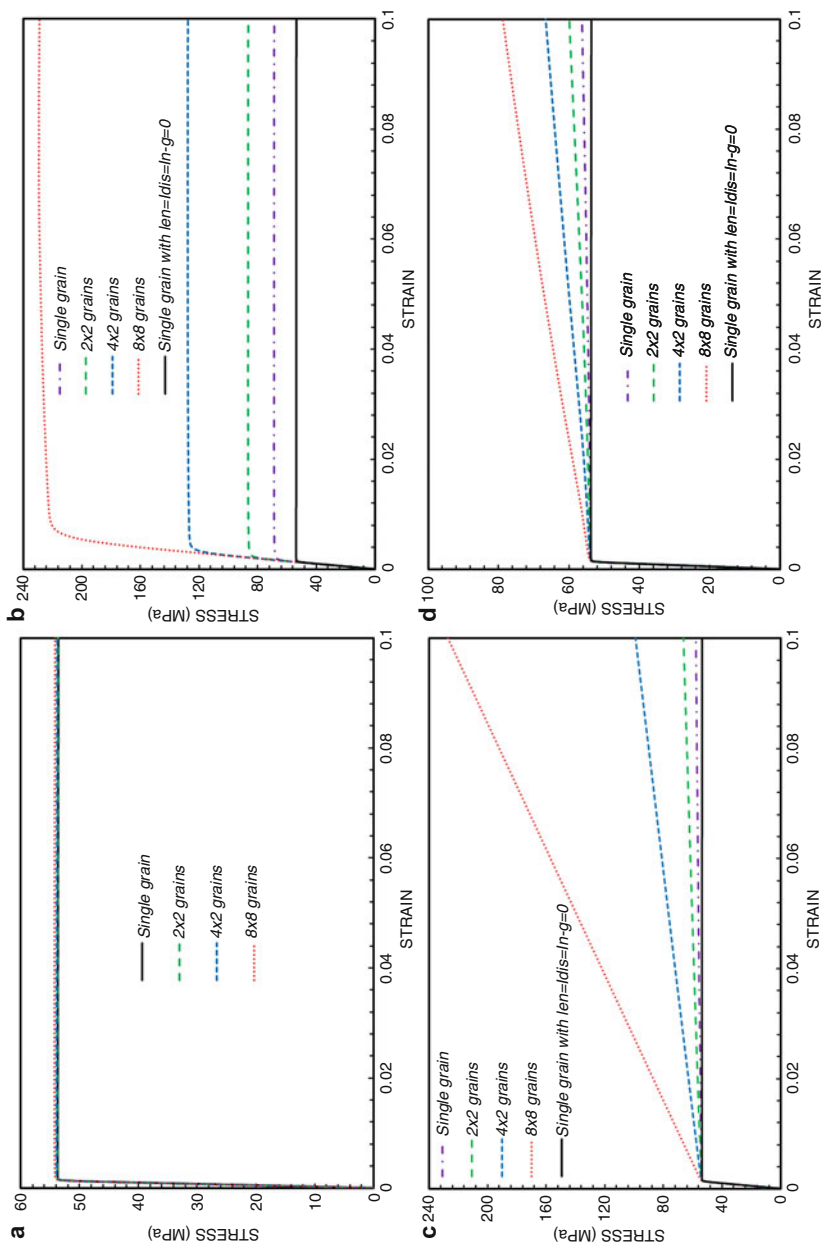


Fig. 13 The effects of the grain boundary area on the stress-strain response with (a) no gradient effect ($\ell_{en} = \ell_{dis} = \ell_{N-G} = 0$), (b) the energetic length scale only ($\ell_{en}/W = 0.1, \ell_{dis} = \ell_{N-G} = 0$), (c) the dissipative length scale only ($\ell_{dis}/W = 0.05, \ell_{en} = \ell_{N-G} = 0$), and (d) the N-G length scale only ($\ell_{N-G}/W = 0.5, \ell_{en} = \ell_{dis} = 0$). (Reprinted with permission from Song and Voyiadjis (2018a))

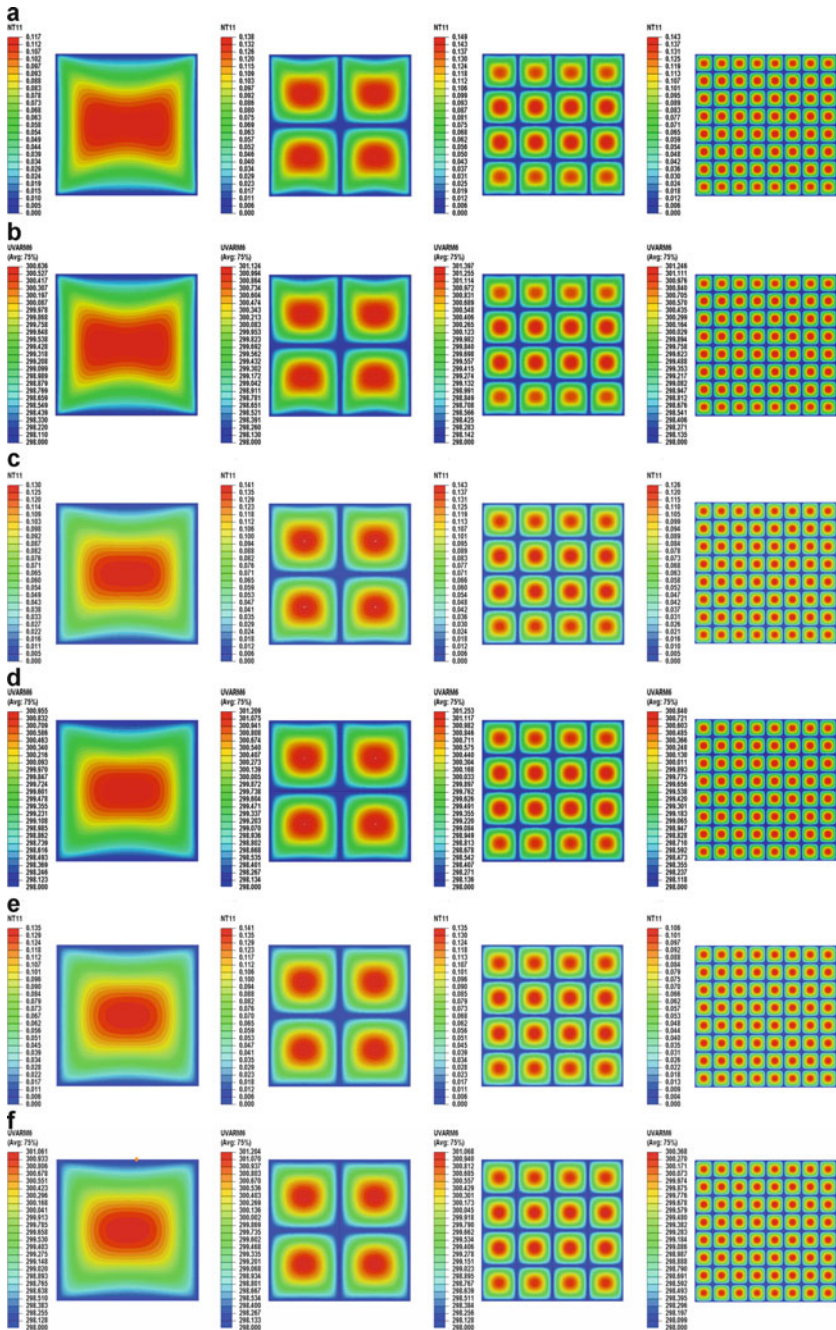


Fig. 14 The distributions of the accumulated plastic strain and the temperature in conjunction with varying grain boundary areas with the energetic length scale only ($\ell_{dis} = \ell_N - G = 0$): (a) e^p with $\ell_{en}/W = 0.1$, (b) T with $\ell_{en}/W = 0.1$, (c) e^p with $\ell_{en}/W = 0.2$, (d) T with $\ell_{en}/W = 0.2$, (e) e^p with $\ell_{en}/W = 0.3$, and (f) T with $\ell_{en}/W = 0.3$. (Reprinted with permission from Song and Voyiadjis (2018a))

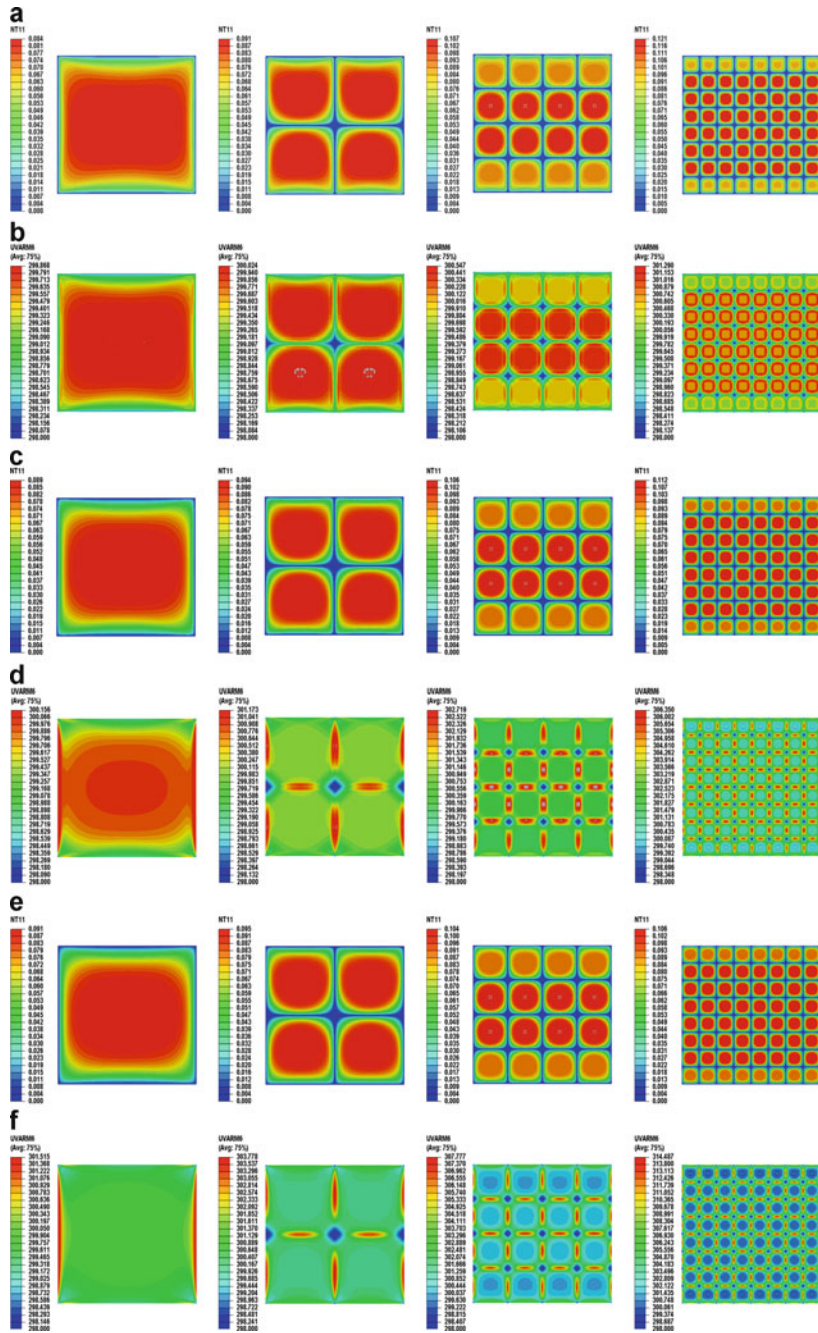


Fig. 15 The distributions of the accumulated plastic strain and the temperature in conjunction with varying grain boundary areas with the dissipative length scale only ($\ell_{en} = \ell_N - G = 0$): (a) e^p with $\ell_{dis}/W = 0.01$, (b) \mathcal{T} with $\ell_{dis}/W = 0.01$, (c) e^p with $\ell_{dis}/W = 0.03$, (d) \mathcal{T} with $\ell_{dis}/W = 0.03$, (e) e^p with $\ell_{dis}/W = 0.05$, and (f) \mathcal{T} with $\ell_{dis}/W = 0.05$. (Reprinted with permission from Song and Voyiadjis (2018a))

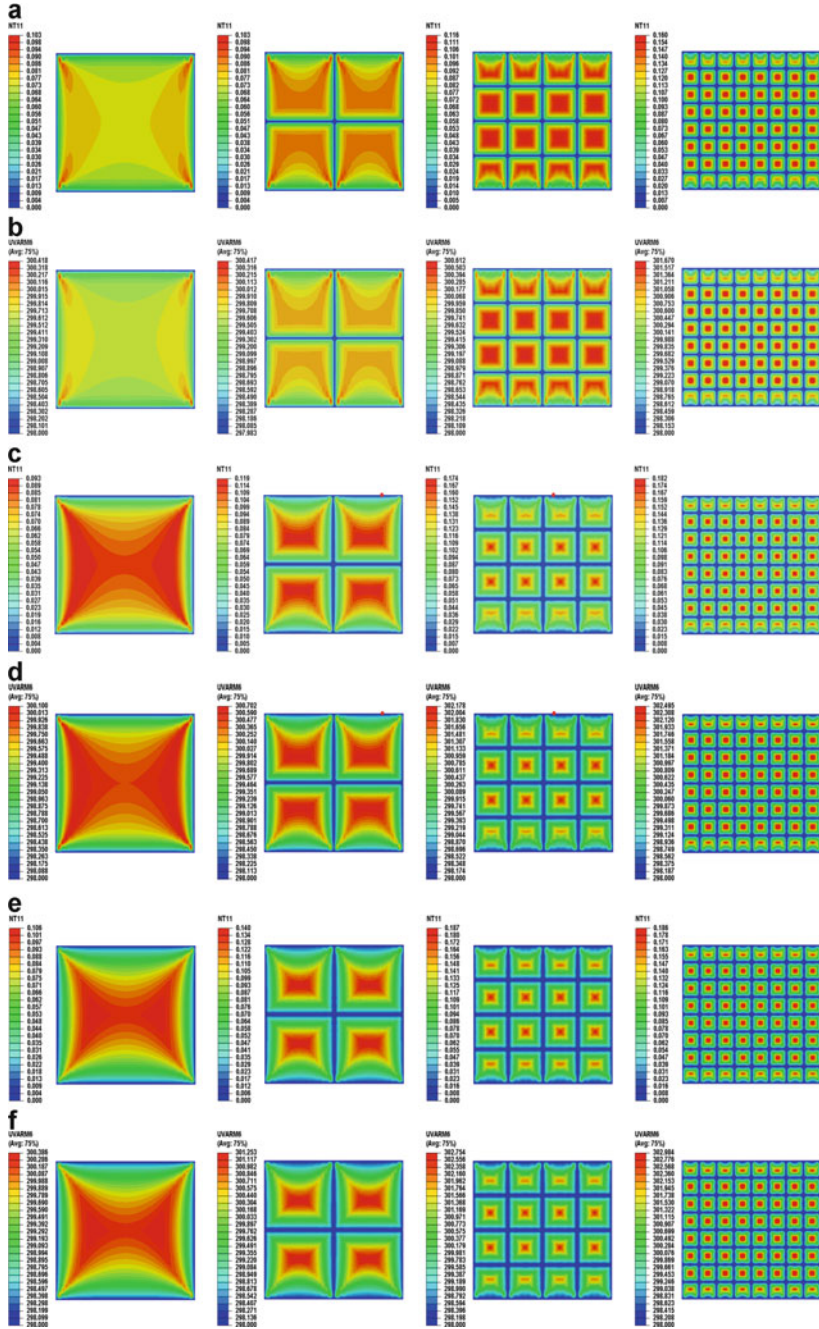


Fig. 16 The distributions of the accumulated plastic strain and the temperature in conjunction with varying grain boundary areas with the N-G length scale only ($\ell_{en} = \ell_{dis} = 0$): (a) e^p with $\ell_{N-G}/W = 0.1$, (b) T with $\ell_{N-G}/W = 0.1$, (c) e^p with $\ell_{N-G}/W = 0.3$, (d) T with $\ell_{N-G}/W = 0.3$, (e) e^p with $\ell_{N-G}/W = 0.5$, and (f) T with $\ell_{N-G}/W = 0$. (Reprinted with permission from Song and Voyiadjis (2018a))

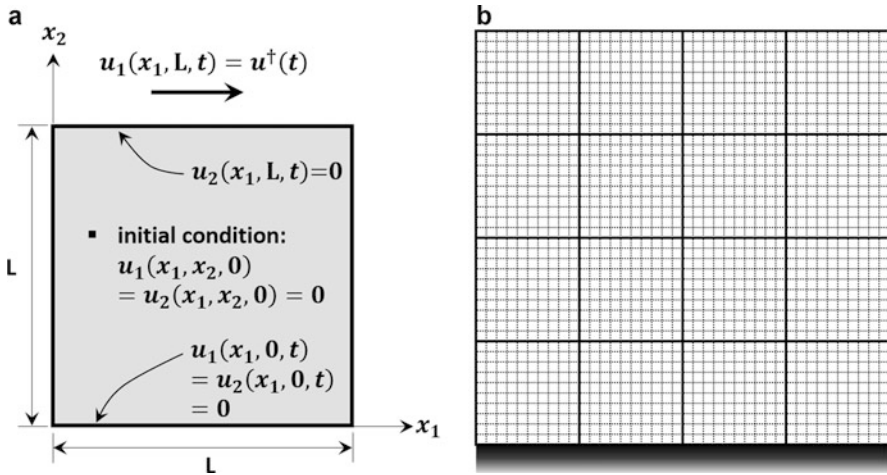


Fig. 17 The schematic illustration of the simple shear problem: (a) the macroscopic, microscopic boundary conditions and initial conditions (b) 4×4 grains. (Reprinted with permission from Song and Voyiadjis (2018b))

boundary flow rule to account for the deformable grain boundary. Furthermore, the characteristics of the proposed strain gradient plasticity theory incorporating the flow rules of both the grain interior and grain boundary are addressed in this section by solving the shear problem of a square plate with an edge of L . The schematic illustration of the problem, the initial conditions, and the macroscopic and microscopic boundary conditions as well as the grain boundary area are shown in Fig. 17. The parameter $u^\dagger(t)$ represents the prescribed displacement. The whole square is meshed using 1,600 (40×40) elements and split into the 16 (4×4) grains by the grain boundary area, which is indicated by the bold lines. The material parameters for the grain interior and grain boundary are presented in Table 3.

As can be seen from Eq. (82), the grain boundary may act like a free surface, i.e., microfree boundary condition, when $\ell_{en}^{GB} = \ell_{dis}^{GB} = 0$. On the other hand, the microhard boundary condition can be compelled under the conditions $\ell_{en}^{GB} \rightarrow \infty$ and $\ell_{dis}^{GB} \rightarrow \infty$. Firstly, the validity of these conditions is examined in this section. Next, the direct comparison between the classical plasticity theory ($\ell_{en}/L = \ell_{dis}/L = \ell_{N-G}/L = 0.0$) and the gradient-enhanced plasticity theory ($\ell_{en}/L = \ell_{dis}/L = \ell_{N-G}/L = 0.1$) is given in order to check the ability of the proposed flow rule on the size effect. The numerical results in terms of the accumulated plastic strain profile and the stress-strain curves are shown in Figs. 18 and 19. As can be seen in these figures, the microfree and microhard boundary conditions are well captured under the classical plasticity theory as well as the gradient-enhanced plasticity theory. In addition, in Fig. 18c, no size effect is observed in the classical plasticity theory with varying normalized material length scales as expected. In Fig. 19c, on the other hand, strain hardening and strengthening are more pronounced as the dimensions of the shear plate height are reduced

Table 3 The material parameters used in section “Intermediate (Deformable) Grain Boundary.” (Reprinted with permission from Song and Voyiadjis (2018b))

Grain interior		Values
E (GPa)	Elastic modulus for isotropic linear elasticity	110
ν	Poisson's ratio	0.33
μ (GPa)	Shear modulus for isotropic linear elasticity	48
ρ ($g \cdot cm^{-3}$)	Density	8.960
c_ϵ (J/g \cdot $^\circ K$)	Specific heat capacity at constant stress	0.385
α^{th} ($\mu m/m \cdot$ $^\circ K$)	Thermal expansion coefficient	16.0
$\dot{\phi}_1, \dot{\phi}_2$ (s^{-1})	Reference plastic strain rate	0.04
r	Nonlinear hardening material constant	0.6
m_1	Nonnegative strain rate sensitivity parameter	0.05
m_2	Nonnegative strain rate sensitivity parameter	0.2
T_y ($^\circ K$)	Thermal material parameter	1,358
n	Temperature sensitivity parameter	0.3
σ_0 (MPa)	Stress-dimensioned scaling constant	195
\mathcal{H}_0 (MPa)	Isotropic hardening parameter	0
Grain boundary		Values
σ_0^{GB} (MPa)	Constant accounting for the GB yield stress	195
\mathcal{H}_0^{GB} (MPa)	GB hardening parameter	0
$\dot{\phi}^{GB}$	Viscous related material parameters for GB	0.04
m^{GB}	Viscous related material parameters for GB	1
T_y^{GB}	GB thermal parameter at the onset of yield	700
n^{GB}	GB thermal parameter	0.4

($\ell_{en}^{GB}/L \rightarrow \infty, \ell_{dis}^{GB}/L \rightarrow \infty$). In Fig. 20, the effects of each material length scale parameter, i.e., ℓ_{en} , ℓ_{dis} , and ℓ_{N-G} , along with the microscopically hard boundary condition, are also examined through the profile of accumulated plastic strain. In addition, the contributions of each length scale parameter on the stress-strain responses are shown in Fig. 20c.

Variations in the stress-strain responses and the evolutions of maximum temperature are investigated for the various values of the normalized energetic and dissipative grain boundary material length scales as shown in Figs. 21 and 22. It is assumed by setting $\ell_{dis}^{GB}/\ell_{dis} = 0$ that all plastic work at the grain boundary is stored as surface energy which depends on the plastic strain state at the surface. In this case, ℓ_{en}^{GB}/ℓ_{en} reflects the grain boundary resistance to plastic deformation. Figures 21b and 22b show the size effects on the strain hardening and temperature evolution due to the grain boundary energetic length scale, and it is more pronounced in the more strongly constrained material, i.e., increasing ℓ_{en}^{GB}/ℓ_{en} . On the other hand, by setting $\ell_{en}^{GB}/\ell_{en} = 0$, it is assumed that the work performed at the grain boundary is dissipated in the absence of surface energy. In this case, $\ell_{dis}^{GB}/\ell_{dis}$ reflects the grain boundary resistance to slip transfer. As can be seen in Fig. 21c, the initial yield strength increases without strain hardening as $\ell_{dis}^{GB}/\ell_{dis}$ increases.

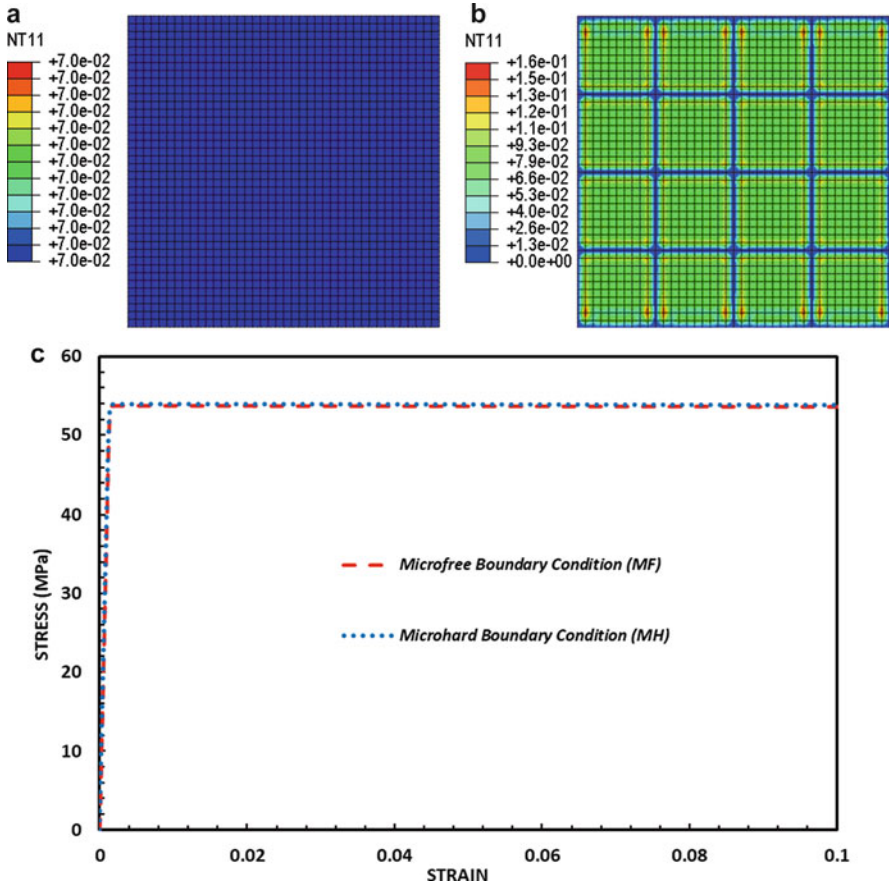


Fig. 18 Classical plasticity theory ($\ell_{en}/L = \ell_{dis}/L = \ell_{N-G}/L = 0.0$). The distributions of the accumulated plastic strain with (a) the microscopically free ($\ell_{en}^{GB} = \ell_{dis}^{GB} = 0$) and (b) microscopically hard boundary conditions ($\ell_{en}^{GB} \rightarrow \infty, \ell_{dis}^{GB} \rightarrow \infty$) and (c) the stress-strain responses. (Reprinted with permission from Song and Voyiadjis (2018b))

Generalized Structure for Modeling Polycrystals from Micro- to Nanoscale Range

In this section, modeling of strengthening in inelastic nanocrystalline materials with reference to the triple junction and grain boundaries using SGP is introduced based on the work of Voyiadjis and Deliktas (2010). Voyiadjis and Deliktas (2010) not only provide the internal interface energies but also introduce two additional internal state variables for the internal surfaces (contact surfaces). By using these internal state variables together with displacement and temperature, the constitutive model is formulated as usual by state laws utilizing free energies and complimentary laws based on the dissipation potentials. One of these new state

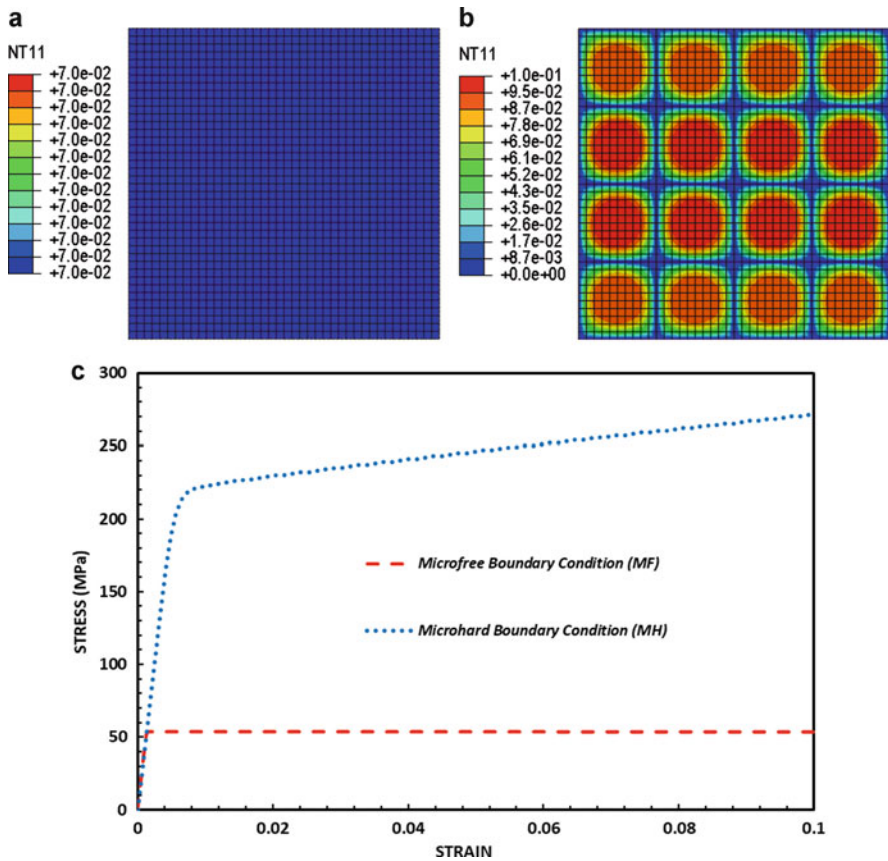


Fig. 19 Strain gradient plasticity theory ($\ell_{en}/L = \ell_{dis}/L = \ell_{N-G}/L = 0.1$). The distributions of the accumulated plastic strain with (a) the microscopically free ($\ell_{en}^{GB} = \ell_{dis}^{GB} = 0$) and (b) microscopically hard boundary conditions ($\ell_{en}^{GB} \rightarrow \infty, \ell_{dis}^{GB} \rightarrow \infty$) and (c) the stress-strain responses. (Reprinted with permission from Song and Voyiadjis (2018b))

variables measures tangential sliding between the grain boundaries and the other measures the respective separation. A homogenization technique is developed to describe the local stress and strain in the material. The material is characterized as a composite with three phases: the grain core, the grain boundaries, and the triple junctions.

Nanocrystalline materials are structurally characterized by a large volume fraction of grain boundaries which may significantly alter their physical, mechanical, and chemical properties in comparison with conventional coarse-grained polycrystalline materials which have grain size usually in the range 10–300 μ m (Meyers et al. 2006). The grain size of nanocrystalline materials is usually below 10 nm. The nanocrystalline structure described by Gleiter (2000) is composed of structural elements with a characteristic size of a few nanometers (see Fig. 23).

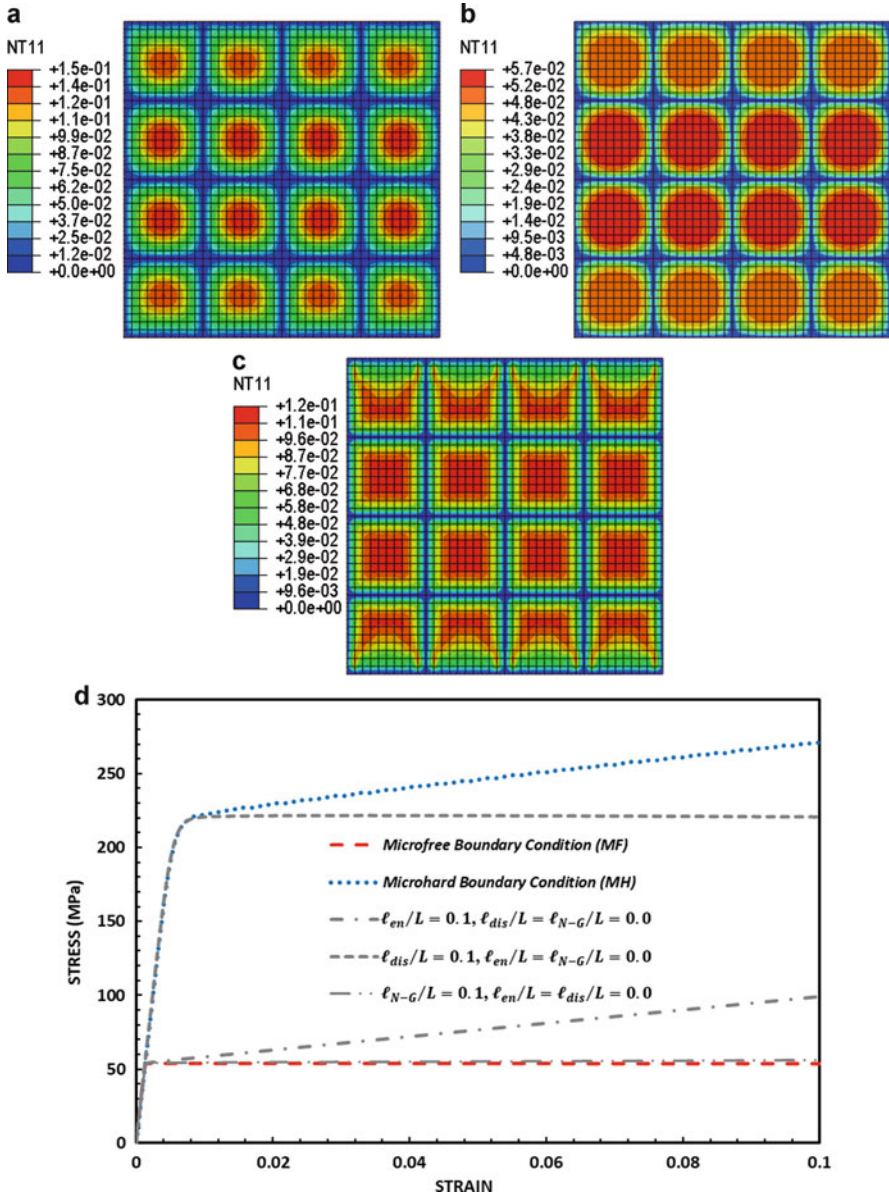


Fig. 20 The distributions of the accumulated plastic strain with the microscopically hard boundary condition ($\ell_{en}^{GB}/L \rightarrow \infty, \ell_{dis}^{GB}/L \rightarrow \infty$) under (a) the energetic length scale only ($\ell_{en}/L = 0.1, \ell_{dis}/L = \ell_{N-G}/L = 0.0$), (b) the dissipative length scale only ($\ell_{dis}/L = 0.1, \ell_{en}/L = \ell_{N-G}/L = 0.0$), (c) the N-G length scale only ($\ell_{N-G}/L = 0.1, \ell_{en}/L = \ell_{dis}/L = 0.0$), and (d) the stress-strain responses. (Reprinted with permission from Song and Voyiadjis (2018b))

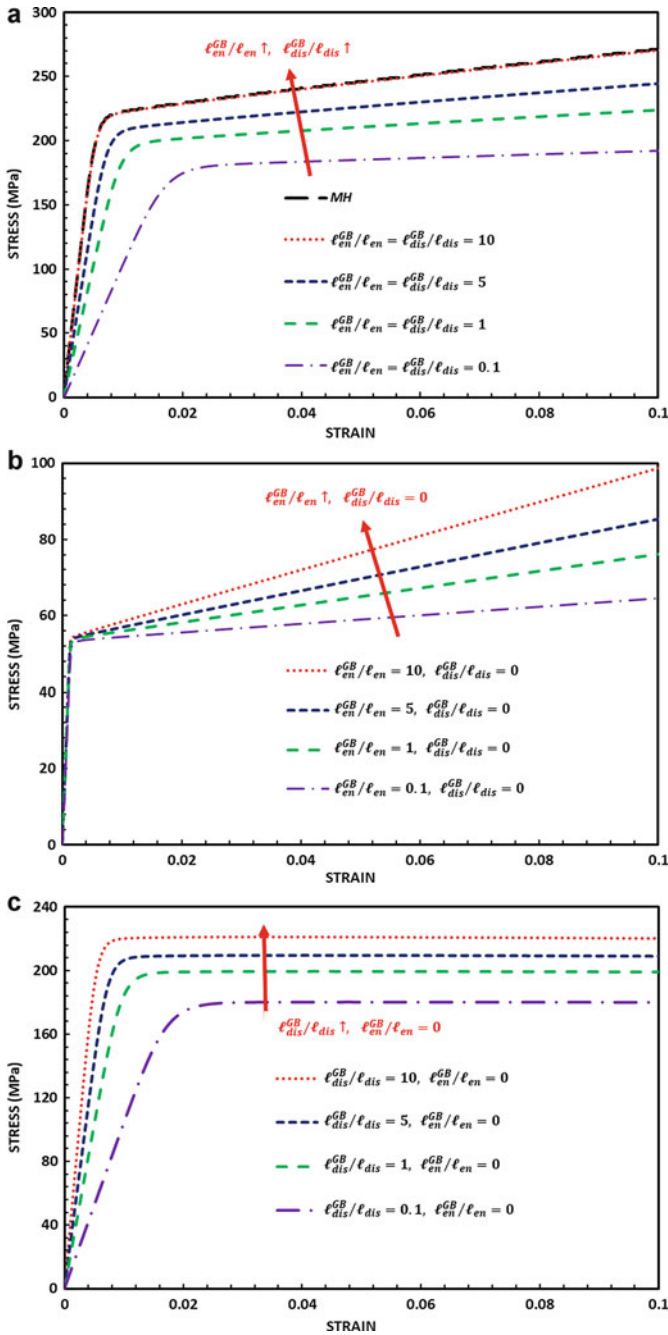


Fig. 21 The distributions of the accumulated plastic strain according to the various values of ℓ_{en}^{GB}/ℓ_{en} and $\ell_{dis}^{GB}/\ell_{dis}$: (a) combined ℓ_{en}^{GB} and ℓ_{dis}^{GB} , (b) ℓ_{en}^{GB} only, and (c) ℓ_{dis}^{GB} only. (Reprinted with permission from Song and Voyiadjis (2018b))

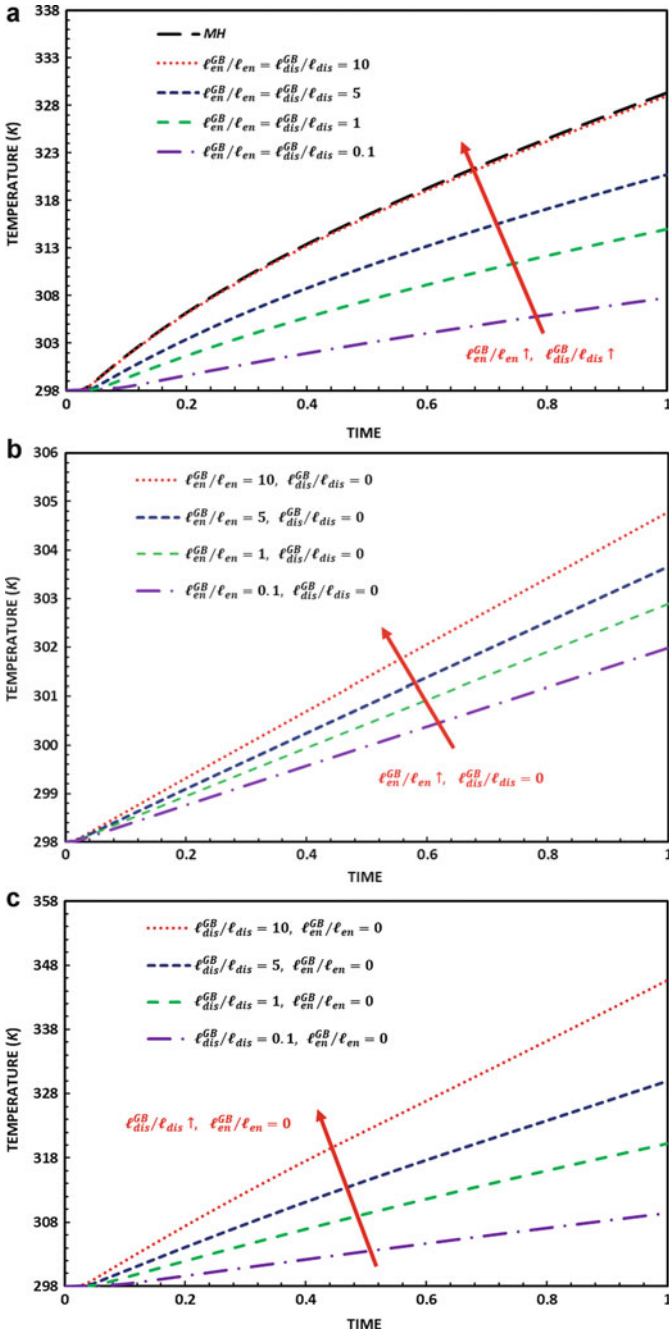


Fig. 22 The evolutions of the maximum temperature according to the various values of ℓ_{en}^{GB}/ℓ_{en} and $\ell_{dis}^{GB}/\ell_{dis}$: (a) combined ℓ_{en}^{GB} and ℓ_{dis}^{GB} , (b) ℓ_{en}^{GB} only, and (c) ℓ_{dis}^{GB} only. (Reprinted with permission from Song and Voyiadjis (2018b))

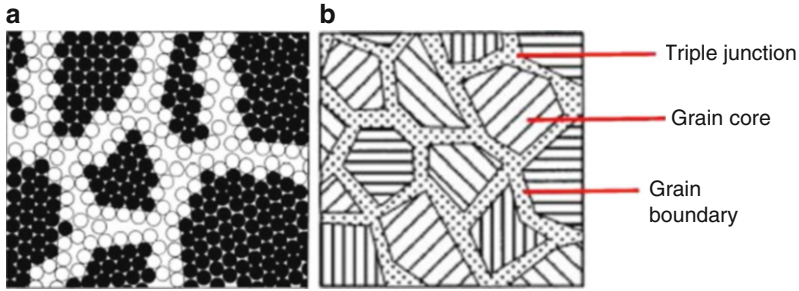


Fig. 23 Two-dimensional model of nanostructured materials (Gleiter 2000; Meyers et al. 2006). (Reprinted with permission from Voyiadjis and Deliktas (2010))

In Fig. 23a, the image represents crystal atoms with neighbor configurations corresponding to the lattice and the boundary atoms with a wide variety of interatomic spacing. The atoms in the centers of the crystals are indicated in black and the one in the boundary regions are represented as open circles. Figure 23b shows the simplified representation of the image of the nanocrystalline structure by comprising grain core, grain boundary, and triple junctions. As one notes from the image in Fig. 23a, many atoms reside in the grain boundary region, and in this case the volume fraction of the interfacial region will not be zero. The small sizes involved in this case limit the conventional operation of the dislocation sources.

The geometrical representation of the RVE proposed by different authors (Mecking and Kocks 1981; Pipard et al. 2009) can be conceptually described by three regions such as the grain core, the grain boundary, and the triple junctions with their corresponding internal interfaces, respectively. In this work the simplified nanocrystalline structure shown in Fig. 23b is represented as a 2D triangle representative volume element (RVE) of a composite material with three phases; grain core, grain boundary, and triple junction (see Fig. 24).

In Fig. 24, if one uses r as the indicator for the phase, then it can be named one of three regions such as the grain core ($r = gc$), the grain boundary ($r = gb$), or the triple junction ($r = tj$). Similarly, one may use the symbol rs to represent the interface surfaces that surround the regions presented in Fig. 24 such that ($rs = cb$) indicates the internal interface between the grain core and grain boundary and ($rs = btj$) shows the internal interface between the grain boundary and the triple junction.

The deformation of each region is attributed to different involved dislocation storage mechanisms and grain boundary sliding or slip. The region called the grain core represents the interior of the grain for which strain hardening only results from the evolution of statistically stored dislocations (SSDs), ρ_s . These are due to multiplication and annihilation processes between dislocations with net Burgers vector equal to zero (Arsenlis and Parks 1999). The SSDs are inherently random within the whole grain. The region with thickness w is termed the grain boundary region and is bounding the grain core. Triple junction is defined as a separate region.

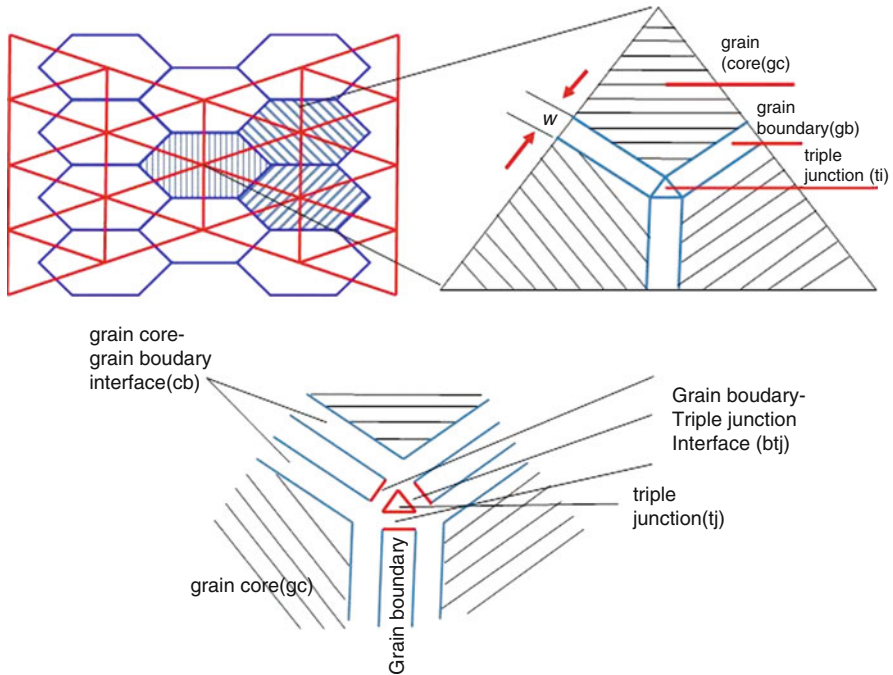


Fig. 24 Representative volume element for the description of the simplified nanocrystalline structure. (Reprinted with permission from Voyiadjis and Deliktas (2010))

The physical origin of the presence of GNDs in this region is due at the mesoscopic scale due to the concentration of slip bands at the grain boundaries such as dislocation pileups. This one has to be relaxed with the help of GNDs. The layer thickness w probably depends on the grain size (since the slip line patterns may vary with grain size) and may vary with strain essentially due to an expansion of the layer with increasing dislocation density. For the sake of simplicity, w will be here assumed to be constant. In addition to these three zones, two interfaces are defined to characterize interactions among the grain core, the grain boundary, and the triple junction. The interface between grain core and the grain boundary that is attributed to the dislocation blocking and pileups can be described as hard, soft, or intermediate interfaces where interfacial hardening and strengthening parameters are introduced based on the theory presented by Voyiadjis and Deliktas (2009b). The second interface which is defined between the grain boundary and the triple junction grain is attributed to the grain boundary sliding or slip which could be described through modeling of the behavior of the triple junction. Full details of the thermodynamical framework for modeling the inelastic behavior of the material and the homogenization scheme for material modeling of the heterogeneous solids developed with a view to address the heterogeneous microstructure and effect of volume fraction of the phases can be found in Voyiadjis and Deliktas (2010).

Multilevel homogenization scheme presented in Voyiadjis and Deliktas (2010) is used for the evaluation of the material model behavior. The local stresses in the grain core, grain boundary, and triple junction are evaluated to better understand the physical mechanisms responsible for the grain size effects on the overall inelastic behavior of the material. The inverse Hall-Petch effects are also investigated, and simulated results are compared with the reported experimental ones. The studied material is Ni-P alloy with a fixed grain boundary width, 4 nm, and varying grain sizes from 10 nm to 200 nm. The material response of each phase is predicted by the constitutive relation given in Voyiadjis and Deliktas (2010). These constitutive relations represent the generalized case of the formulation. For example, if one considers the polycrystalline structured material, the inelastic deformation is mostly dominated in the grain core by dislocation mechanisms (movement and storage), and both plastic strain and its gradient are used in the formulation characterized by the strain gradient formulation. In the case of nanocrystalline structured materials, for the grain size range (Pipard et al. 2009), $30 \text{ nm} < d \leq 100 \text{ nm}$, it is assumed that the grain core carries only statistically stored dislocations, whereas the grain boundary and the triple junction regions have geometrically necessary dislocations. Therefore, in this case, while the plastic strain gradient is dropping out from the grain core phase, it should be considered in the grain boundary and triple junction phases. In the last case where nanocrystalline grain size is less than 30 nm (Meyers et al. 2006), the dislocation-based inelastic deformation in the grain core phase is essentially shut off, and the inelastic deformation occurs primarily by slip and separation of the grain boundaries. The material constants of the Ni-P are obtained from experimental studies (Benson et al. 2001; Zhao et al. 2003) reported by Qing and Xingming (2006).

The overall stress-strain curves of the Ni-P nanocrystalline material for various grain sizes are presented in Fig. 25. Clearly materials with large grain sizes show more plastic hardening than that with smaller grain sizes. For material grain size that is bigger, the volume fraction of the interphases is negligible, and therefore, the material response is dominated by the behavior of the grain core. However, when the grain size is getting smaller, the overall response is more complex. In this case the material responses of both the grain boundary and the triple junction along with that of the grain core play essential role in the overall behavior of the nanocrystalline material.

Nanocrystalline material exhibits almost an order of magnitude higher strain rate sensitivity than their microcrystalline counterparts. This enhanced strain rate sensitivity of the nanocrystalline materials is captured qualitatively by proposed theory here (Fig. 26).

As one notes clearly from Fig. 26, the nanocrystalline material shows the strong strain rate sensitivity at the grain size of 20 nm. However, as the grain size is increased that of microcrystalline counterpart it is observed that material does not exhibit the strain rate sensitivity.

Finally, in Fig. 27, the relation between the yield strength and the grain size in nanometer Ni-P is shown. By comparison with experimental results, when the critical size of the crystalline grains of nanometer Ni-P is greater than 10 nm, the

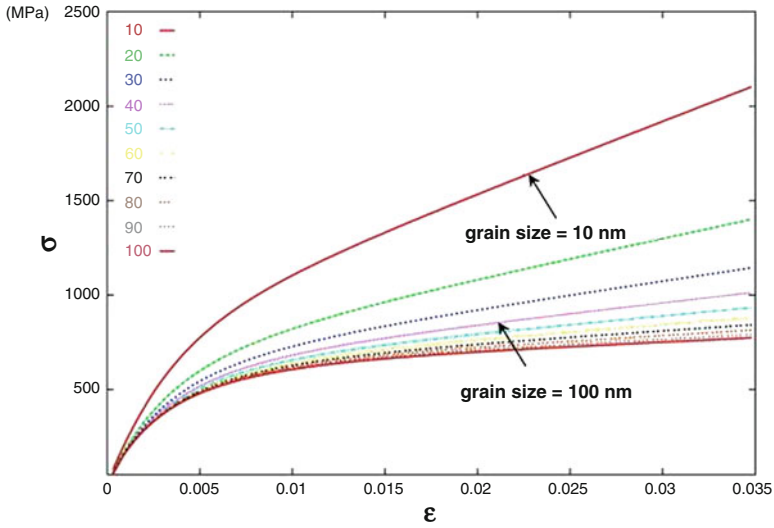


Fig. 25 Overall stress-strain response of nanocrystalline with various grain sizes. (Reprinted with permission from Voyiadjis and Deliktas (2010))

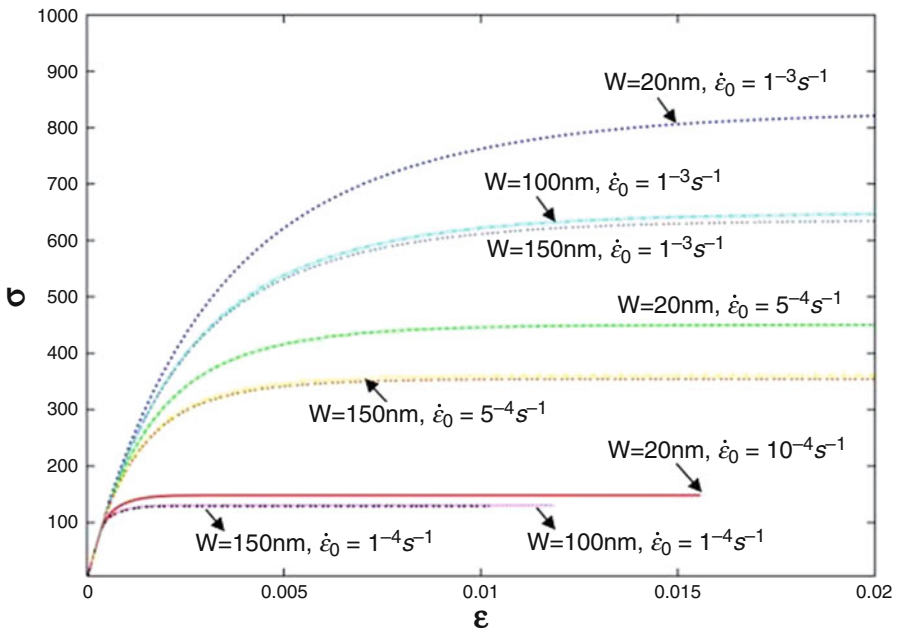


Fig. 26 Stress-strain curves for the nanocrystalline with the grain sizes 20 nm, 100 nm, and 150 nm at the different strain rates. (Reprinted with permission from Voyiadjis and Deliktas (2010))

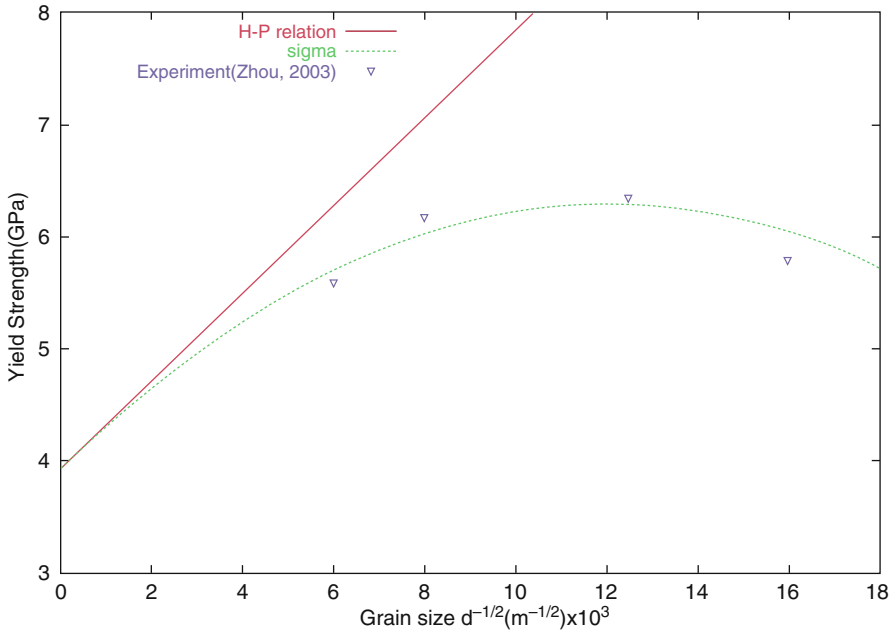


Fig. 27 Grain size dependency of yield strength for the Ni-P material. (Reprinted with permission from Voyiadjis and Deliktas (2010))

positive H-P slope can be seen. However, when the size of the crystalline grain is less than 10 nm, a negative H-P slope can be observed (Fig. 27).

Conclusions

The two-dimensional finite element analysis for the thermodynamically consistent thermomechanical coupled gradient-enhanced plasticity model is proposed within the areas of grain interior and grain boundary and validated by comparing against two sets of small-scale experiments demonstrating the size effects. The proposed formulation is developed based on the concept of dislocation interaction mechanisms and thermal activation energy. The thermodynamic microstresses are assumed to be divided in two components, i.e., the energetic and dissipative components, which in turn, both energetic and dissipative material length scale parameters are incorporated in the governing constitutive equations. These two thermodynamic microstresses can be respectively obtained in a direct way from the Helmholtz free energy and rate of dissipation potential by taking maximum entropy production into account. The concept of GND density is additionally employed in this work to interpret the microstructural strengthening mechanisms induced by the nonhomogeneous deformation. Correspondingly, the model in this work

incorporates the terms related to GND-induced strengthening and the additional material length scale parameter. Not only the partial heat dissipation due to fast transient time but also the distribution of the temperature caused by the transition from plastic work to heat is included into the coupled thermomechanical model by deriving a generalized heat equation.

The proposed SGP model and the correspondingly developed finite element code are validated through the comparisons against the experimental measurements from small-scale aluminum and nickel thin film tests. The material parameters for two kinds of metals are also calibrated by using the experimental measurements. The numerical results show good agreement with the experimental measurements in terms of both tests. The simple shear problem and the square plate problem are solved based on the validated model in order to examine size effect in small-scale metallic materials and two null boundary conditions, respectively. The energetic hardening, dissipative strengthening, and GND hardening are well observed, respectively, with varying ℓ_{en} , ℓ_{dis} , and ℓ_{N-G} . The parametric study is carried out to investigate the effect of the temperature-related parameters on the stress-strain curve, the distributions of the temperature and the accumulated plastic strain, and the temperature evolution. The effect of the microhard boundary condition at the grain boundary on the stress-strain curve and the distributions of the plastic strain and the temperature are presented by solving the square plate problem. The strengthening effect due to the microhard boundary condition at the grain boundary is clearly observed with increasing grain boundary areas in this simulation. In addition, the microfree and microhard boundary conditions are well captured by using the proposed grain boundary flow rule. Lastly, the size effects on the stress-strain responses and the evolutions of maximum temperature are well observed with the cases of (a) combined ℓ_{en}^{GB} and ℓ_{dis}^{GB} , (b) ℓ_{en}^{GB} only, and (c) ℓ_{dis}^{GB} only.

A generalized structure for modeling polycrystals from micro- to nano-size range is presented in the last section. The polycrystalline structure is defined in terms of the grain core, the grain boundary, and the triple junction regions with their corresponding volume fractions. This is achieved by describing a simplified nanocrystalline structure as a 2D triangle representative volume element (RVE) of a composite material. It is shown that the developed model is able to capture qualitatively the enhanced strain rate sensitivity of the nanocrystalline materials that exhibit almost an order of magnitude higher strain rate sensitivity than their microcrystalline counterparts. It is also shown that the inverse Hall-Petch is captured from the model prediction that fits quite well to the experimental observations.

Appendix A. Deriving the Balance Equations

The total strain rate $\dot{\varepsilon}_{ij}$ is defined as follows:

$$\dot{\varepsilon}_{ij} = \frac{1}{2} (\dot{u}_{i,j} + \dot{u}_{j,i}) \quad (103)$$

with the velocity gradient $\dot{u}_{i,j} = \partial \dot{u}_i / \partial x_j$.

By substituting Eq. (2) into Eq. (8), one obtains

$$\mathcal{P}^{int} = \int_{\Omega_0} \left(\sigma_{ij} \dot{\epsilon}_{ij} - \sigma_{ij} \dot{\epsilon}_{ij}^p + \mathbf{x} \dot{\epsilon}^p + Q_i \dot{\epsilon}_{i,i}^p + \mathcal{A} \dot{\mathcal{T}} + \mathcal{B}_i \dot{\mathcal{T}}_{i,i} \right) dV \tag{104}$$

From the plastic incompressibility ($\dot{\epsilon}_{kk}^p = 0$), $\sigma_{ij} \dot{\epsilon}_{ij}^p = \bar{\sigma}_{ij} \dot{\epsilon}_{ij}^p$. The divergence theorem can be used here in Eq. (104) together with Eq. (103) to obtain the following expression:

$$\begin{aligned} \mathcal{P}^{int} = & \int_{\partial\Omega_0} (\sigma_{ij} n_j \dot{u}_i + Q_i n_i \dot{\epsilon}^p + \mathcal{B}_i n_i) dS \\ & - \int_{\Omega_0} \left(\sigma_{ij,j} \dot{u}_i - \bar{\sigma}_{ij} \dot{\epsilon}_{ij}^p + \mathbf{x} \dot{\epsilon}^p - Q_{i,i} \dot{\epsilon}^p + \mathcal{A} \dot{\mathcal{T}} - \mathcal{B}_{i,i} \dot{\mathcal{T}} \right) dV \end{aligned} \tag{105}$$

By equating the external power given in Eq. (9) to the internal power ($\mathcal{P}^{int} = \mathcal{P}^{ext}$), the following expression is obtained:

$$\begin{aligned} & \int_{\partial\Omega_0} \{ (\sigma_{ij} n_j - \mathfrak{t}_i) \dot{u}_i + (Q_i n_i - \mathfrak{m}) \dot{\epsilon}^p + (\mathcal{B}_i n_i - \mathfrak{a}) \} dS \\ & - \int_{\Omega_0} \{ (\sigma_{ij,j} + \mathfrak{l}_i) \dot{u}_i + (\bar{\sigma}_{ij} N_{ij} - \mathbf{x} + Q_{i,i}) \dot{\epsilon}^p + (\mathcal{A} - \mathcal{B}_{i,i}) \dot{\mathcal{T}} \} dV \end{aligned} \tag{106}$$

Here, \dot{u}_i , $\dot{\epsilon}^p$, and $\dot{\mathcal{T}}$ can be designated randomly when the following conditions are satisfied:

$$\sigma_{ij,j} + \mathfrak{l}_i = 0 \tag{107}$$

$$\bar{\sigma}_{ij} = (\mathbf{x} - Q_{k,k}) N_{ij} \tag{108}$$

$$\mathcal{B}_{i,i} - \mathcal{A} = 0 \tag{109}$$

$$\mathfrak{t}_j = \sigma_{ij} n_i \tag{110}$$

$$\mathfrak{m} = Q_i n_i \tag{111}$$

$$\mathfrak{a} = \mathcal{B}_i n_i \tag{112}$$

References

- Abaqus, *User's Manual (Version 6.12)* (Dassault Systemes Simulia Corporation, Providence, 2012)
- E.C. Aifantis, *J. Eng. Mater-T Asme.* **106**, 4 (1984)
- E.C. Aifantis, *Int. J. Plast.* **3**, 3 (1987)
- K.E. Aifantis, J.R. Willis, *J. Mech. Phys. Solids* **53**, 5 (2005)
- A.H. Almasri, G.Z. Voyiadjis, *Acta Mech.* **209**, 1–2 (2010)
- L. Anand, M.E. Gurtin, S.P. Lele, C. Gething, *J. Mech. Phys. Solids* **53**, 8 (2005)
- A. Arsenlis, D.M. Parks, *Acta Mater.* **47**, 5 (1999)
- L. Bardella, *J. Mech. Phys. Solids* **54**, 1 (2006)
- D.J. Benson, H.H. Fu, M.A. Meyers, *Mater. Sci. Eng. A* **319**, 319–321 (2001)
- T. Bergeman, J. Qi, D. Wang, Y. Huang, H.K. Pechkis, E.E. Eyler, P.L. Gould, W.C. Stwalley, R.A. Cline, J.D. Miller, D.J. Heinzen, *J. Phys. B-Atom. Mol. Opt. Phys.* **39**, 19 (2006)
- S.D. Brorson, A. Kazeroonian, J.S. Moodera, D.W. Face, T.K. Cheng, E.P. Ippen, M.S. Dresselhaus, G. Dresselhaus, *Phys. Rev. Lett.* **64**, 18 (1990)
- H.B. Callen, *Thermodynamics and an Introduction to Thermostatistics* (Wiley, Hoboken, 1985)
- P. Cermelli, M.E. Gurtin, *Int. J. Solids Struct.* **39**, 26 (2002)
- W.A.T. Clark, R.H. Wagoner, Z.Y. Shen, T.C. Lee, I.M. Robertson, H.K. Birnbaum, *Scr. Met. Mater.* **26**, 2 (1992)
- B.D. Coleman, W. Noll, *Arch. Ration. Mech. Anal.* **13**, 3 (1963)
- D. Faghihi, G.Z. Voyiadjis, *J. Eng. Mater-T Asme.* **136**, 1 (2014)
- N.A. Fleck, J.W. Hutchinson, *Adv. Appl. Mech.* **33**, 295 (1997)
- N.A. Fleck, J.W. Hutchinson, *J. Mech. Phys. Solids* **49**, 10 (2001)
- N.A. Fleck, G.M. Muller, M.F. Ashby, J.W. Hutchinson, *Acta Metall. Mater.* **42**, 2 (1994)
- N.A. Fleck, J.W. Hutchinson, J.R. Willis, *P. Roy. Soc. A-Math. Phys.* **470**, 2170 (2014)
- N.A. Fleck, J.W. Hutchinson, J.R. Willis, *J. Appl. Mech-T Asme.* **82**, 7 (2015)
- S. Forest, *J. Eng. Mech-Asce.* **135**, 3 (2009)
- S. Forest, M. Amestoy, *Cr. Mecanique* **336**, 4 (2008)
- P. Fredriksson, P. Gudmundson, *Model. Simul. Mater. Sci.* **15**, 1 (2007)
- A. Garroni, G. Leoni, M. Ponsiglione, *J. Eur. Math. Soc.* **12**, 5 (2010)
- S. Giacomazzi, F. Leroy, C. L'Henaff, J.J. Joffraud, *Lett. Appl. Microbiol.* **38**, 2 (2004)
- H. Gleiter, *Acta Mater.* **48**, 1 (2000)
- P. Gudmundson, *J. Mech. Phys. Solids* **52**, 6 (2004)
- M.E. Gurtin, *Int. J. Plast.* **19**, 1 (2003)
- M.E. Gurtin, *J. Mech. Phys. Solids* **52**, 11 (2004)
- M.E. Gurtin, *J. Mech. Phys. Solids* **56**, 2 (2008)
- M.E. Gurtin, L. Anand, *J. Mech. Phys. Solids* **53**, 7 (2005)
- M.E. Gurtin, L. Anand, *J. Mech. Phys. Solids* **57**, 3 (2009)
- M.E. Gurtin, E. Fried, L. Anand, *The Mechanics and Thermodynamics of Continua* (Cambridge University Press, Cambridge, 2010)
- S. Han, T. Kim, H. Lee, H. Lee, *Electronics System-Integration Technology Conference, 2008. ESTC 2008. 2nd* (2008)
- M.A. Haque, M.T.A. Saif, *Acta Mater.* **51**, 11 (2003)
- J.P. Hirth, J. Lothe, *Theory of Dislocations* (Krieger Publishing Company, 1982). ISBN: 0894646176, 9780894646171
- Y. Huang, J. Qi, H.K. Pechkis, D. Wang, E.E. Eyler, P.L. Gould, W.C. Stwalley, *J. Phys. B-Atom. Mol. Opt. Phys.* **39**, 19 (2006)
- J.W. Hutchinson, *Acta Mech. Sinica-Prc.* **28**, 4 (2012)
- J.S. Hwang, H.L. Park, T.W. Kim, H.J. Lee, *Phys. Status Solidi a-Appl. Res.* **148**, 2 (1995)
- A.M. Ivanitsky, D.A. Kadakov, *Izv. Vyssh. Uchebn. Zaved. Radiofiz.* **26**, 9 (1983)
- T.W. Kim, H.L. Park, *J. Cryst. Growth* **159**, 1–4 (1996)

- T.W. Kim, H.L. Park, J.Y. Lee, *Appl. Phys. Lett.* **64**, 19 (1994)
- T.C. Lee, I.M. Robertson, H.K. Birnbaum, *Scr. Metall.* **23**, 5 (1989)
- S.P. Lele, L. Anand, *Philos. Mag.* **88**, 30–32 (2008)
- J.M. Lim, K. Cho, M. Cho, *Appl. Phys. Lett.* **110**, 1 (2017)
- W.J. Liu, K. Saanouni, S. Forest, P. Hu, *J. Non-Equil. Thermody.* **42**, 4 (2017)
- V.A. Lubarda, *Int. J. Solids Struct.* **45**, 1 (2008)
- D.L. McDowell, *Int. J. Plast.* **26**, 9 (2010)
- H. Mecking, U.F. Kocks, *Acta Metall.* **29**, 1865–1875 (1981)
- M.A. Meyers, A. Mishra, D.J. Benson, *Prog. Mater. Sci.* **51**, 427 (2006)
- H.B. Muhlhaus, E.C. Aifantis, *Int. J. Solids Struct.* **28**, 7 (1991)
- L. Nicola, Y. Xiang, J.J. Vlassak, E. Van der Giessen, A. Needleman, *J. Mech. Phys. Solids* **54**, 10 (2006)
- W.D. Nix, H.J. Gao, *J. Mech. Phys. Solids* **46**, 3 (1998)
- T. Ohmura, A.M. Minor, E.A. Stach, J.W. Morris, *J. Mater. Res.* **19**, 12 (2004)
- N. Ohno, D. Okumura, *J. Mech. Phys. Solids* **55**, 9 (2007)
- P.A. Parilla, M.F. Hundley, A. Zettl, *Solid State Commun.* **87**, 6 (1993)
- H.L. Park, S.H. Lee, T.W. Kim, *Compd. Semicond.* **1995**, 145 (1996)
- J.M. Pipard, N. Nicaise, S. Berbenni, O. Bouaziz, M. Berveiller, *Comput. Mater. Sci.* **45**, 604–610 (2009)
- X. Qing, G. Xingming, *Int. J. Solids Struct.* **43**, 25–26 (2006)
- W.A. Soer, K.E. Aifantis, J.T.M. De Hosson, *Acta Mater.* **53**, 17 (2005)
- Y. Song, G.Z. Voyiadjis, *Int. J. Solids Struct.* **134**, 195–215 (2018a)
- Y. Song, G.Z. Voyiadjis, *J. Theor. App. Mech-Pol.* **56**, 2 (2018b)
- S. Sun, B.L. Adams, W.E. King, *Philos. Mag. A* **80**, 1 (2000)
- D.Y. Tzou, Y.S. Zhang, *Int. J. Eng. Sci.* **33**, 10 (1995)
- R. Venkatraman, P.R. Besser, J.C. Bravman, S. Brennan, *J. Mater. Res.* **9**, 2 (1994)
- E. Voce, *Meta* **51**, 219 (1955)
- G.Z. Voyiadjis, B. Deliktas, *Int. J. Plast.* **25**, 10 (2009a)
- G.Z. Voyiadjis, B. Deliktas, *Int. J. Eng. Sci.* **47**, 11–12 (2009b)
- G.Z. Voyiadjis, B. Deliktas, *Acta Mech.* **213**, 1–2 (2010)
- G.Z. Voyiadjis, D. Faghihi, *Int. J. Plast.* **30–31**, 218 (2012)
- G.Z. Voyiadjis, R. Peters, *Acta Mech.* **211**, 1–2 (2010)
- G.Z. Voyiadjis, Y. Song, *Philos. Mag.* **97**, 5 (2017)
- G.Z. Voyiadjis, C. Zhang, *Mat. Sci. Eng. A-Struct.* **621**, 218 (2015)
- G.Z. Voyiadjis, A.H. Almasri, T. Park, *Mech. Res. Commun.* **37**, 3 (2010)
- G.Z. Voyiadjis, D. Faghihi, Y.D. Zhang, *Int. J. Solids Struct.* **51**, 10 (2014)
- G.Z. Voyiadjis, Y. Song, T. Park, *J. Eng. Mater. Technol.* **139**, 2 (2017)
- Y. Xiang, J.J. Vlassak, *Acta Mater.* **54**, 20 (2006)
- M. Yaghoobi, G.Z. Voyiadjis, *Acta Mater.* **121**, 190 (2016)
- C. Zhang, G.Z. Voyiadjis, *Mat. Sci. Eng. A-Struct.* **659**, 55 (2016)
- B. Zhang, Y. Song, G.Z. Voyiadjis, W.J. Meng, *J. Mater. Res.* **361**(1–2), 160–164 (2018)
- J.R. Zhao, J.C. Li, Q. Jiang, *J. Alloys Compd.* **361**, 160 (2003)



Fractional Nonlocal Continuum Mechanics and Microstructural Models

23

Vasily E. Tarasov

Contents

Introduction	840
Long-Range Interactions of Lattice Particles	841
Lattices Fractional Integro-Differentiation	842
Lattice Derivatives of Integer Orders	843
From Lattice Models to Continuum Models	844
Lattice Derivatives Are Exact Discretization of Continuum Derivatives	845
References	847

Abstract

Models of physical lattices with long-range interactions for nonlocal continuum are suggested. The lattice long-range interactions are described by exact fractional-order difference operators. Continuous limit of suggested lattice operators gives continuum fractional derivatives of non-integer orders. The proposed approach gives a new microstructural basis to formulation of theory of nonlocal materials with power-law nonlocality. Moreover these lattice models, which is based on exact fractional differences, allow us to have a unified microscopic description of fractional nonlocal and standard local continuum.

V. E. Tarasov (✉)

Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, Moscow, Russia
e-mail: tarasov@theory.sinp.msu.ru

© Springer Nature Switzerland AG 2019

G. Z. Voyiadjis (ed.), *Handbook of Nonlocal Continuum Mechanics for Materials and Structures*, https://doi.org/10.1007/978-3-319-58729-5_15

839

Keywords

Non-local continuum · Lattice model · Long-range interaction · Fractional derivatives · Exact differences

Introduction

Continuum mechanics can be considered as a continuous limit of lattice models, where the length scales of a continuum element are much larger than the distances between the lattice particles. In general, continuum models, which are described by differential equations with a finite number of integer-order derivatives with respect to coordinates, cannot be considered as nonlocal models. An application of general form infinite series with integer derivatives to describe media with nonlocality is a difficult problem. This problem can be solved for power-law type of nonlocality by using fractional derivatives of non-integer orders. It is important to note that the use of the derivatives of non-integer orders is actually equivalent to using an infinite number of derivatives of integer orders (e.g., see Lemma 15.3 in Samko et al. 1993). The fractional-order derivatives allow us to describe continuum with nonlocality of power-law type. First time the fractional derivatives with respect to space coordinates have been applied to mechanics of nonlocal continuum by Gubenko (1957) and Rostovtsev (1959) in 1957. Recently, the fractional-order derivatives are actively used to describe continua with power-law type of nonlocality in Di Paola et al. (2009a,b, 2013), Di Paola and Zingales (2011), Drapaca and Sivaloganathan (2012), Challamel et al. (2013), Atanackovic et al. (2014a,b), Tarasov (2010, 2013, 2014a,b,c,d, 2015a,b,c,d,e,g, 2016b,c,e, 2017), Carpinteri et al. (2009, 2011), Sapora et al. (2013), Cottone et al. (2009a,b), Sumelka and Blaszczyk (2014), Sumelka et al. (2015), and Sumelka (2015), where the microscopic models of fractional continuum mechanics are also discussed. Fractional calculus is a powerful tool to describe processes in continuously distributed media with nonlocality of power-law type. As it was shown in Tarasov (2010, 2006a,b), the continuum equations with fractional derivatives are directly connected to lattice models with long-range interactions. As it was shown in Tarasov (2006a,b), the differential equations with fractional derivatives of non-integer orders can be derived from equation for lattice particles with long-range interactions in the continuous limit, where the distance between the lattice particles tends to zero. A direct connection between the lattice with long-range interaction and nonlocal continuum has been proved by using the special transform operation (Tarasov 2006a,b, 2014e, 2015f) (see also Tarasov and Zaslavsky 2006a,b).

In works Tarasov (2014e, 2015f,h, 2016f,g,h,i, 2017b), it has been suggested exact lattice (discrete) analogs of fractional differential operators of integer and non-integer orders. This mathematical tool allows us to formulate lattice models that are exact discrete (microstructural) analogs of continuum models. The models of lattices with long-range interactions and corresponding models of continua with power-law nonlocality have been suggested in Tarasov (2013, 2014a,b,c,d, 2015a,b,c,e,d, 2016b,c,e, 2017).

Long-Range Interactions of Lattice Particles

Let us consider three-dimensional physical lattices. These lattices are characterized by space periodicity. For unbounded lattices we can use three noncoplanar vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$, which are the shortest vectors by which a lattice can be displaced to coincidence with itself. Sites of this lattice can be characterized by the number vector $\mathbf{n} = (n_1, n_2, n_3)$, where n_i ($j = 1, 2, 3$) are integer. For simplification, we consider a lattice with mutually perpendicular primitive lattice vectors \mathbf{a}_j , ($j = 1, 2, 3$). This means that we use a primitive orthorhombic Bravais lattice. We choose directions of the axes of the Cartesian coordinate system coincide with the vector \mathbf{a}_j , such that $\mathbf{a}_j = a_j \mathbf{e}_j$, where $a_i = |\mathbf{a}_j| > 0$ and \mathbf{e}_j are the basis vectors of the Cartesian coordinate system. Then the vector \mathbf{n} can be represented as $\mathbf{n} = n_1 \mathbf{e}_1 + n_2 \mathbf{e}_2 + n_3 \mathbf{e}_3$.

Choosing a coordinate origin at one of the lattice sites, then the positions of all other site with $\mathbf{n} = (n_1, n_2, n_3)$ is described by the vector $\mathbf{r}(\mathbf{n}) = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3$. The lattice sites are numbered by \mathbf{n} . Therefore the vector \mathbf{n} is called the number vector of the corresponding particle. We assume that the equilibrium positions of particles coincide with the lattice sites $\mathbf{r}(\mathbf{n})$. Coordinates $\mathbf{r}(\mathbf{n})$ of lattice sites differ from the coordinates of the corresponding particles, when particles are displaced with respect to their equilibrium positions. To define the coordinates of a particle, we define displacement of this particle from its equilibrium position by the vector field $\mathbf{u}(\mathbf{n}, t) = \sum_{i=1}^3 u_i(\mathbf{n}, t) \mathbf{e}_i$, where $u_i(\mathbf{n}, t) = u_i(n_1, n_2, n_3, t)$ are components of the displacement vector for lattice particle.

To simplify our consideration, we use some assumptions for microscopic (lattice) structure. We assume that lattice n -particle interacts by pair manner with all lattice m -particles that reflects the long-range nature of the interaction in the suggested nonlocal model of material.

In general, it is possible to consider the long-range interactions that are characterized by different orders α_j in different directions $\mathbf{e}_j = \mathbf{a}_j / |\mathbf{a}_j|$. In these models we should use the difference operators of orders α_j , where $\alpha_1 \neq \alpha_2, \alpha_1 \neq \alpha_3, \alpha_2 \neq \alpha_3$.

Let us give a definition of the long-range interaction of power-law type (for details see Tarasov (2006a,b) and Sect. 8 of Tarasov 2010). An interaction of lattice particles is called the interaction of power-law type if the kernels $K(n - m)$ of this interaction satisfy the conditions

$$\lim_{k \rightarrow 0} \frac{\hat{K}_\alpha(k) - \hat{K}_\alpha(0)}{|k|^\alpha} = A_\alpha, \quad \alpha > 0, \quad 0 < |A_\alpha| < \infty, \quad (1)$$

where

$$\hat{K}_\alpha(k \Delta x) = \sum_{n=-\infty}^{+\infty} e^{-ikn\Delta x} K(n) = 2 \sum_{n=1}^{\infty} K(n) \cos(kn\Delta x). \quad (2)$$

The examples of power-law type interaction are considered in Tarasov (2006a,b, 2010, 2014c,e).

The power-law type of interactions is suggested in the papers Tarasov (2006a,b) (see also Sect. 8 of Tarasov 2010, 2014e, 2015f). This type of interactions is characterized by the power-law asymptotic behavior of spatial dispersions in lattice. We assume that the power-law spatial dispersion in the lattice can be caused by non-Debye screening of electromagnetic interatomic interactions. Some aspects of the theory of this screening are described in the papers Tarasov and Trujillo (2013) and Tarasov (2016a), where fractional-order power-law spatial dispersion in electrodynamics of continuum is discussed. Some elasticity models of materials with power-law spatial dispersion are discussed in Tarasov (2013, 2014a).

Lattices Fractional Integro-Differentiation

In the lattice models, the long-range interactions can be described by using the lattice fractional integro-differential operators that has been suggested in Tarasov (2014e, 2015f, 2016f,g).

Let us give a definition of the lattice fractional integro-differentiation: *Lattice integro-differential operators* ${}^T\mathbb{D}_L \left[\begin{smallmatrix} \alpha \\ j \end{smallmatrix} \right]$ of order $\alpha > -1$ in the direction $\mathbf{e}_j = \mathbf{a}_j / |\mathbf{a}_j|$ is an operator that is defined by the equation

$$\left({}^T\mathbb{D}_L \left[\begin{smallmatrix} \alpha \\ j \end{smallmatrix} \right] f \right) (\mathbf{n}) = \frac{1}{a_j^\alpha} \sum_{m_j=-\infty}^{+\infty} K_\alpha(n_j - m_j) f(\mathbf{m}), \quad (j = 1, \dots, N), \quad (3)$$

where $\mathbf{n}, \mathbf{m} \in \mathbb{Z}^N$, and $\mathbf{n} = \mathbf{m} + (n_j - m_j)\mathbf{e}_j$. The kernel $K_\alpha(n)$ is a real-valued function of integer variable $n \in \mathbb{Z}$, that is defined by the equation

$$K_\alpha(n) = \cos\left(\frac{\pi \alpha}{2}\right) \frac{\pi^\alpha}{\alpha + 1} {}_1F_2\left(\frac{\alpha + 1}{2}; \frac{1}{2}, \frac{\alpha + 3}{2}; -\frac{\pi^2 n^2}{4}\right) - \sin\left(\frac{\pi \alpha}{2}\right) \frac{\pi^{\alpha+1} n}{\alpha + 2} {}_1F_2\left(\frac{\alpha + 2}{2}; \frac{3}{2}, \frac{\alpha + 4}{2}; -\frac{\pi^2 n^2}{4}\right), \quad (4)$$

where $\alpha > -1$. If $\alpha > 0$, then the lattice operator (3) will be called the lattice fractional derivative. For $-1 < \alpha < 0$, then the lattice operator (3) is called the lattice fractional integral (antiderivative).

In kernel (4) of the lattice derivatives, we use the generalized hypergeometric function ${}_1F_2$ that is defined (see Sect. 1.6 of Kilbas et al. 2006) by the equation

$${}_1F_2(a; b, c; z) := \sum_{k=0}^{\infty} \frac{(a)_k}{(b)_k (c)_k} \frac{z^k}{k!} = \sum_{k=0}^{\infty} \frac{\Gamma(a+k) \Gamma(b) \Gamma(c)}{\Gamma(a) \Gamma(b+k) \Gamma(c+k)} \frac{z^k}{k!}, \quad (5)$$

where $(a)_k$ is the Pochhammer symbol (rising factorial) that is defined by $(a)_0 = 1$, and $(a)_k = \Gamma(a + k)/\Gamma(a)$.

Note that using (5) and $\Gamma(z + 1) = z\Gamma(z)$, expression (4) of the kernel $K_\alpha(n)$ can be represented in the form

$$K_\alpha(n) = \sum_{k=0}^{\infty} \frac{(-1)^k \pi^{2k+1/2+\alpha} n^{2k}}{2^{2k} k! \Gamma(k + 1/2)} \left(\frac{\cos\left(\frac{\pi\alpha}{2}\right)}{\alpha + 2k + 1} - \frac{\pi n \sin\left(\frac{\pi\alpha}{2}\right)}{(\alpha + 2k + 2)(2k + 1)} \right), \tag{6}$$

where $\alpha > -1$.

Lattice Derivatives of Integer Orders

Let us give exact expressions of lattice fractional derivatives for integer orders (Tarasov 2014e, 2015h,f, 2016f). The lattice derivatives of integer orders are defined by the equations

$$\mathcal{T} \Delta^{2s} f[n] := \sum_{\substack{m=-\infty \\ m \neq 0}}^{+\infty} K_{2s}(m) f[n - m] + K_{2s}(0) f[n], \quad (s \in \mathbb{N}), \tag{7}$$

$$\mathcal{T} \Delta^{2s-1} f[n] := \sum_{\substack{m=-\infty \\ m \neq 0}}^{+\infty} K_{2s-1}(m) f[n - m], \quad (s \in \mathbb{N}), \tag{8}$$

where the kernels $K_{2s}(m)$ and $K_{2s-1}(m)$ are defined by Eqs. (3) and (6).

The kernels of these lattice operators can be represented (Tarasov 2016f) in a simpler form.

$$K_{2s}(m) = \sum_{k=0}^{s-1} \frac{(-1)^{m+k+s} (2s)! \pi^{2s-2k-2}}{(2s - 2k - 1)!} \frac{1}{m^{2k+2}} \quad (m \in \mathbb{Z}, \quad m \neq 0), \tag{9}$$

$$K_{2s-1}(m) = \sum_{k=0}^{s-1} \frac{(-1)^{m+k+s+1} (2s - 1)! \pi^{2s-2k-2}}{(2s - 2k - 1)!} \frac{1}{m^{2k+1}} \quad (m \in \mathbb{Z}, \quad m \neq 0). \tag{10}$$

and

$$K_{2s}(0) = \frac{(-1)^s \pi^{2s}}{2s + 1}, \quad K_{2s-1}(0) = 0. \tag{11}$$

Note that $K_{2s}(0)$ describes a self-interaction of lattice particles. The interaction of different particles is described by $K_s(n - m)$ with $n - m \neq 0$.

The first-order lattice operators is

$$\mathcal{T} \Delta^1 f[n] := \sum_{\substack{m=-\infty \\ m \neq 0}}^{+\infty} \frac{(-1)^m}{m} f[n-m]. \quad (12)$$

The lattice antiderivative is defined by the equation

$$\mathcal{T} \Delta^{-1} f[n] := \sum_{\substack{m=-\infty \\ m \neq 0}}^{+\infty} \pi^{-1} Si(\pi m) f[n-m], \quad (13)$$

where $Si(z)$ is the sine integral.

From Lattice Models to Continuum Models

Using the methods suggested in Tarasov (2006a,b, 2014e, 2015f, 2016f), we can define the operation that transforms a lattice field $u(\mathbf{n})$ into a field $u(\mathbf{r})$ of continuum. For this transformation, we will consider the lattice scalar field $u(\mathbf{n})$ as Fourier series coefficients of some function $\hat{u}(\mathbf{k})$ for $k_j \in [-k_{j0}/2, k_{j0}/2]$, where $j = 1, 2, 3$. At the next step, we use the continuous limit $\mathbf{k}_0 \rightarrow \infty$ to obtain $\tilde{u}(\mathbf{k})$. Finally we apply the inverse Fourier integral transformation to obtain the continuum scalar field $u(\mathbf{r})$. Let us describe these steps with details:

Step 1: The discrete Fourier series transform $u(\mathbf{n}) \rightarrow \mathcal{F}_\Delta\{u(\mathbf{n})\} = \hat{u}(\mathbf{k})$ of the lattice scalar field $u(\mathbf{n})$ is defined by

$$\hat{u}(\mathbf{k}) = \mathcal{F}_\Delta\{u(\mathbf{n})\} = \sum_{n_1, n_2, n_3=-\infty}^{+\infty} u(\mathbf{n}) e^{-i(\mathbf{k}, \mathbf{r}(\mathbf{n}))}, \quad (14)$$

where $\mathbf{r}(\mathbf{n}) = \sum_{j=1}^3 n_j \mathbf{a}_j$, and $a_j = 2\pi/k_{j0}$ are distance between lattice particles in the direction \mathbf{a}_j .

Step 2: The passage to the limit $\hat{u}(\mathbf{k}) \rightarrow \text{Lim}\{\hat{u}(\mathbf{k})\} = \tilde{u}(\mathbf{k})$, where we use $a_j \rightarrow 0$ (or $k_{j0} \rightarrow \infty$), allows us to derive the function $\tilde{u}(\mathbf{k})$ from $\hat{u}(\mathbf{k})$. By definition $\tilde{u}(\mathbf{k})$ is the Fourier integral transform of the continuum field $u(\mathbf{r})$, and the function $\hat{u}(k)$ is the Fourier series transform of the lattice field $u(\mathbf{n})$, where

$$u(\mathbf{n}) = \frac{(2\pi)^3}{k_{10}k_{20}k_{30}} u(\mathbf{r}(\mathbf{n})),$$

and $\mathbf{r}(\mathbf{n}) = \sum_{j=1}^3 n_j \mathbf{a}_j = \sum_{j=1}^3 2\pi n_j / k_{j0} \rightarrow \mathbf{r}$.

Step 3: The inverse Fourier integral transform $\tilde{u}(\mathbf{k}) \rightarrow \mathcal{F}^{-1}\{\tilde{u}(\mathbf{k})\} = u(\mathbf{r})$ is defined by

$$u(\mathbf{r}) = \frac{1}{(2\pi)^3} \iiint_{-\infty}^{+\infty} dk_1 dh_2 dk_3 e^{i \sum_{j=1}^3 k_j x_j} \tilde{u}(\mathbf{k}) = \mathcal{F}^{-1}\{\tilde{u}(\mathbf{k})\}. \tag{15}$$

The combination $\mathcal{F}^{-1} \circ \text{Lim} \circ \mathcal{F}_\Delta$ of the operations \mathcal{F}^{-1} , Lim , and \mathcal{F}_Δ define the lattice-continuum transform operation

$$\mathcal{T}_{L \rightarrow C} = \mathcal{F}^{-1} \circ \text{Lim} \circ \mathcal{F}_\Delta \tag{16}$$

that maps lattice models into the continuum models (Tarasov 2006a,b).

The lattice-continuum transform operation $\mathcal{T}_{L \rightarrow C}$ as the combination of three operations $\mathcal{F}^{-1} \circ \text{Lim} \circ \mathcal{F}_\Delta$ can be applied not only for lattice fields but also for lattice operators. The operation $\mathcal{T}_{L \rightarrow C}$ allows us to map of lattice derivatives ${}^T\mathbb{D}_L \left[\begin{smallmatrix} \alpha \\ j \end{smallmatrix} \right]$ into continuum derivatives ${}^{RT}\mathbb{D}_C \left[\begin{smallmatrix} \alpha \\ j \end{smallmatrix} \right]$, by the equation

$${}^{RT}\mathbb{D}_C \left[\begin{smallmatrix} \alpha \\ j \end{smallmatrix} \right] f(\mathbf{r}) = \mathcal{T}_{L \rightarrow C} \left(\left({}^T\mathbb{D}_L \left[\begin{smallmatrix} \alpha \\ j \end{smallmatrix} \right] f \right) (\mathbf{n}) \right) \quad (\alpha > -1). \tag{17}$$

The transform operation $\mathcal{T}_{L \rightarrow C}$, which maps the lattice operator to the continuum operators, has been defined in Tarasov (2006a,b, 2010, 2016f).

For integer order $\alpha > 0$, the suggested lattice fractional derivatives are directly related to the partial derivatives of integer orders

$${}^{RT}\mathbb{D}_C \left[\begin{smallmatrix} m \\ j \end{smallmatrix} \right] f(\mathbf{r}) = \mathcal{T}_{L \rightarrow C} \left(\left({}^T\mathbb{D}_L \left[\begin{smallmatrix} m \\ j \end{smallmatrix} \right] f \right) (\mathbf{n}) \right) = \frac{\partial^m f(\mathbf{r})}{\partial x_j^m}, \tag{18}$$

where $m \in \mathbb{N}$. This means that the lattice fractional derivatives (3) of integer positive values of $\alpha = m \in \mathbb{N}$ are standard partial derivatives of integer orders m .

Lattice Derivatives Are Exact Discretization of Continuum Derivatives

In paper Tarasov (2016f), it has been suggested an exact discretization for derivatives of integer and non-integer orders.

Let us give the definition of the exact discretization of integro-differentiation: *Let $\mathcal{A}(\mathbb{R}^N)$ be a function space and ${}^{RT}\mathbb{D}_C \left[\begin{smallmatrix} \alpha \\ j \end{smallmatrix} \right]$ be an differential operator on a function space $\mathcal{A}(\mathbb{R}^N)$ such that*

$${}^{RT}\mathbb{D}_C \left[\begin{smallmatrix} \alpha \\ j \end{smallmatrix} \right] f(\mathbf{r}) = g(\mathbf{r}) \quad (\mathbf{r} \in \mathbb{R}^N) \tag{19}$$

for all $f(\mathbf{r}) \in \mathcal{A}(\mathbb{R}^N)$, where $g(\mathbf{r}) \in \mathcal{A}(\mathbb{R}^N)$. The lattice operator ${}^T\mathbb{D}_L \begin{bmatrix} \alpha \\ j \end{bmatrix}$ will be called an exact discretization of ${}^{RT}\mathbb{D}_C \begin{bmatrix} \alpha \\ j \end{bmatrix}$ if the equation

$${}^T\mathbb{D}_L \begin{bmatrix} \alpha \\ j \end{bmatrix} f(\mathbf{n}) = g(\mathbf{n}) \quad (\mathbf{n} \in \mathbb{Z}^N) \tag{20}$$

is satisfied for all $f(\mathbf{n}) \in \mathcal{A}(\mathbb{Z}^N)$, where $g(\mathbf{n}) \in \mathcal{A}(\mathbb{Z}^N)$.

Condition (20) means that the equality

$$\left({}^{RT}\mathbb{D}_C \begin{bmatrix} \alpha \\ j \end{bmatrix} f(\mathbf{r}) \right)_{\mathbf{r}=\mathbf{n}} = {}^T\mathbb{D}_L \begin{bmatrix} \alpha \\ j \end{bmatrix} f(\mathbf{n}) \tag{21}$$

holds for all $\mathbf{n} \in \mathbb{Z}^N$.

In work Tarasov (2016f), it has been proposed a general principle of algebraic correspondence for an exact discretization. For the lattice fractional calculus, this principle can be formulated in the following form.

Principle of exact correspondence between lattice and continuum theories:

The correspondence between the theories of difference (lattice) and differential (integro-differential) equations lies not so much in the limiting condition when the primitive lattice vectors $a_j \rightarrow 0$ as in the fact that mathematical operations on these two theories should obey in many cases the same laws.

This principle is similar to the Dirac’s principle of correspondence between the quantum and classical theories. To specify our consideration, the suggested principle of correspondence can be formulated in the following form: *The lattice integro-differential operators, which are exact discretization of continuum integro-differential operators of integer or non-integer orders, should satisfy the same algebraic characteristic relations as the continuum integro-differential operators.* For detail see Tarasov (2016f).

Let us give the necessary condition of exact discretization of derivatives: *The lattice operator (difference) ${}^T\mathbb{D}_L \begin{bmatrix} \alpha_j \\ j \end{bmatrix}$ is the exact discretization of the continuum operator ${}^{RT}\mathbb{D}_C \begin{bmatrix} \alpha_j \\ j \end{bmatrix}$ of order α_j if the following condition*

$$\frac{1}{a_j} \mathcal{F}^{-1} \mathcal{F}_{a,\Delta} \left({}^T\mathbb{D}_L \begin{bmatrix} \alpha_j \\ j \end{bmatrix} \right) = {}^{RT}\mathbb{D}_C \begin{bmatrix} \alpha_j \\ j \end{bmatrix} \tag{22}$$

is satisfied for arbitrary values of $a_j > 0$. If the condition (22) is not satisfied, but the condition

$$\lim_{a_j \rightarrow 0^+} \frac{1}{a_j} \mathcal{F}^{-1} \mathcal{F}_{a,\Delta} \left({}^T\mathbb{D}_L \begin{bmatrix} \alpha_j \\ j \end{bmatrix} \right) = {}^{RT}\mathbb{D}_C \begin{bmatrix} \alpha_j \\ j \end{bmatrix} \tag{23}$$

holds, then the lattice operator ${}^T\mathbb{D}_L \left[\begin{smallmatrix} \alpha_j \\ j \end{smallmatrix} \right]$ is the asymptotic (approximate) discretization of the continuum operator ${}^R\mathbb{D}_C \left[\begin{smallmatrix} \alpha_j \\ j \end{smallmatrix} \right]$. Here \mathcal{F}^{-1} is the inverse Fourier integral transform and $\mathcal{F}_{a,\Delta}$ is the Fourier series transform. It is obvious that the lattice operator, which is an exact discretization, satisfies the condition (23) of the asymptotic discretization also.

Using the suggested lattice derivatives and antiderivative, we can obtain exact lattice analogs of the differential equations of continuum models without using approximations. It should be noted that this discretization allows to obtain lattice (difference) equations whose solutions are equal to the solutions of corresponding differential equations (Tarasov 2016f,h, 2017a). The suggested lattice fractional derivatives greatly simplify the construction of microstructural (lattice) models with long-range interactions in fractional nonlocal theories of continua, media, and fields (Tarasov 2013, 2014a,b,c,d, 2015a,b,c, 2016b,c).

References

- T. Atanackovic, S. Pilipovic, B. Stankovic, D. Zorica, *Fractional Calculus with Applications in Mechanics: Vibrations and Diffusion Processes* (Wiley-ISTE, Hoboken, 2014a)
- T.M. Atanackovic, S. Pilipovic, B. Stankovic, D. Zorica, *Fractional Calculus with Applications in Mechanics: Wave Propagation, Impact and Variational Principles* (Wiley-ISTE, Hoboken, 2014b)
- A. Carpinteri, P. Cornetti, A. Sapora, Static-kinematic fractional operators for fractal and non-local solids. *Zeitschrift für Angewandte Mathematik und Mechanik. Appl. Math. Mech.* **89**(3), 207–217 (2009)
- A. Carpinteri, P. Cornetti, A. Sapora, A fractional calculus approach to nonlocal elasticity. *Eur. Phys. J. Spec. Top.* **193**, 193–204 (2011)
- N. Challamel, D. Zorica, T.M. Atanackovic, D.T. Spasic, On the fractional generalization of Eringen's nonlocal elasticity for wave propagation. *C. R. Mec.* **341**(3), 298–303 (2013)
- G. Cottone, M. Di Paola, M. Zingales, Elastic waves propagation in 1D fractional non-local continuum. *Physica E* **42**(2), 95–103 (2009a)
- G. Cottone, M. Di Paola, M. Zingales, Fractional mechanical model for the dynamics of non-local continuum, in *Advances in Numerical Methods*. Lecture Notes in Electrical Engineering, vol. 11 (Springer, New York, 2009b), Chapter 33. pp. 389–423
- M. Di Paola, M. Zingales, Fractional differential calculus for 3D mechanically based non-local elasticity. *Int. J. Multiscale Comput. Eng.* **9**(5), 579–597 (2011)
- M. Di Paola, F. Marino, M. Zingales, A generalized model of elastic foundation based on long-range interactions: integral and fractional model. *Int. J. Solids Struct.* **46**(17), 3124–3137 (2009a)
- M. Di Paola, G. Failla, M. Zingales, Physically-based approach to the mechanics of strong non-local linear elasticity theory. *J. Elast.* **97**(2), 103–130 (2009b)
- M. Di Paola, G. Failla, A. Pirrotta, A. Sofi, M. Zingales, The mechanically based non-local elasticity: an overview of main results and future challenges. *Philos. Trans. R. Soc. A.* **371**(1993), 20120433 (2013)
- C.S. Drapaca, S. Sivaloganathan, A fractional model of continuum mechanics. *J. Elast.* **107**(2), 105–123 (2012)
- V.S. Gubenko, Some contact problems of the theory of elasticity and fractional differentiation. *J. Appl. Math. Mech.* **21**(2), 279–280 (1957, in Russian)

- A.A. Kilbas, H.M. Srivastava, J.J. Trujillo, *Theory and Applications of Fractional Differential Equations* (Elsevier, Amsterdam, 2006) p. 353
- N.A. Rostovtsev, Remarks on the paper by V.S. Gubenko, Some contact problems of the theory of elasticity and fractional differentiation. *J. Appl. Math. Mech.* **23**(4), 1143–1149 (1959)
- S.G. Samko, A.A. Kilbas, O.I. Marichev, *Fractional Integrals and Derivatives Theory and Applications* (Gordon and Breach, New York, 1993), p. 1006
- A. Saporita, P. Cornetti, A. Carpinteri, Wave propagation in nonlocal elastic continua modelled by a fractional calculus approach. *Commun. Nonlinear Sci. Numer. Simul.* **18**(1), 63–74 (2013)
- W. Sumelka, Non-local KirchhoffLove plates in terms of fractional calculus. *Arch. Civil Mech. Eng.* **15**(1), 231–242 (2015)
- W. Sumelka, T. Blaszczyk, Fractional continua for linear elasticity. *Arch. Mech.* **66**(3), 147–172 (2014)
- W. Sumelka, R. Zaera, J. Fernández-Sáez, A theoretical analysis of the free axial vibration of non-local rods with fractional continuum mechanics. *Meccanica.* **50**(9), 2309–2323 (2015)
- V.E. Tarasov, Continuous limit of discrete systems with long-range interaction. *J. Phys. A.* **39**(48), 14895–14910 (2006a). arXiv:0711.0826
- V.E. Tarasov, Map of discrete system into continuous. *J. Math. Phys.* **47**(9), 092901 (2006b). arXiv:0711.2612
- V.E. Tarasov, *Fractional Dynamics: Applications of Fractional Calculus to Dynamics of Particles, Fields and Media* (Springer, New York, 2010)
- V.E. Tarasov, Lattice model with power-law spatial dispersion for fractional elasticity. *Centr. Eur. J. Phys.* **11**(11), 1580–1588 (2013)
- V.E. Tarasov, Fractional gradient elasticity from spatial dispersion law. *ISRN Condens. Matter Phys.* **2014**, 794097 (13 pages) (2014a)
- V.E. Tarasov, Lattice model of fractional gradient and integral elasticity: long-range interaction of Grunwald-Letnikov-Riesz type. *Mech. Mater.* **70**(1), 106–114 (2014b). arXiv:1502.06268
- V.E. Tarasov, Lattice with long-range interaction of power-law type for fractional non-local elasticity. *Int. J. Solids Struct.* **51**(15–16), 2900–2907 (2014c). arXiv:1502.05492
- V.E. Tarasov, Fractional quantum field theory: from lattice to continuum. *Adv. High Energy Phys.* **2014**, 957863 (14 pages) (2014d)
- V.E. Tarasov, Toward lattice fractional vector calculus. *J. Phys. A.* **47**(35), 355204 (51 pages) (2014e)
- V.E. Tarasov, General lattice model of gradient elasticity. *Mod. Phys. Lett. B.* **28**(7), 1450054 (2014f). arXiv:1501.01435
- V.E. Tarasov, Three-dimensional lattice approach to fractional generalization of continuum gradient elasticity. *Prog. Frac. Differ. Appl.* **1**(4), 243–258 (2015a)
- V.E. Tarasov, Fractional-order difference equations for physical lattices and some applications. *J. Math. Phys.* **56**(10), 103506 (2015b)
- V.E. Tarasov, Discretely and continuously distributed dynamical systems with fractional nonlocality, in *Fractional Dynamics*, ed. by C. Cattani, H.M. Srivastava, X.-J. Yang (De Gruyter Open, Berlin, 2015c), Chapter 3, pp. 31–49. <https://doi.org/10.1515/9783110472097-003>
- V.E. Tarasov, Variational principle of stationary action for fractional nonlocal media. *Pac. J. Math. Ind.* **7**(1), Article 6. [11 pages] (2015d)
- V.E. Tarasov, Non-linear fractional field equations: weak non-linearity at power-law non-locality. *Nonlinear Dyn.* **80**(4), 1665–1672 (2015e)
- V.E. Tarasov, Lattice fractional calculus. *Appl. Math. Comput.* **257**, 12–33 (2015f)
- V.E. Tarasov, Lattice model with nearest-neighbor and next-nearest-neighbor interactions for gradient elasticity. *Discontinuity Nonlinearity Complex* **4**(1), 11–23 (2015g). arXiv:1503.03633
- V.E. Tarasov, Exact discrete analogs of derivatives of integer orders: differences as infinite series. *J. Math.* **2015**, Article ID 134842 (2015h)
- V.E. Tarasov, Electric field in media with power-law spatial dispersion. *Mod. Phys. Lett. B* **30**(10), 1650132 (11 pages) (2016a). <https://doi.org/10.1142/S0217984916501323>
- V.E. Tarasov, Discrete model of dislocations in fractional nonlocal elasticity. *J. King Saud Univ. Sci.* **28**(1), 33–36 (2016b)

- V.E. Tarasov, Three-dimensional lattice models with long-range interactions of Grunwald-Letnikov type for fractional generalization of gradient elasticity. *Meccanica*. **51**(1), 125–138 (2016c)
- V.E. Tarasov, Fractional mechanics of elastic solids: continuum aspects. *J. Eng. Mech.* **143**(5), (2017). [https://doi.org/10.1061/\(ASCE\)EM.1943-7889.0001074](https://doi.org/10.1061/(ASCE)EM.1943-7889.0001074)
- V.E. Tarasov, Partial fractional derivatives of Riesz type and nonlinear fractional differential equations. *Nonlinear Dyn.* **86**(3), 1745–1759 (2016e). <https://doi.org/10.1007/s11071-016-2991-y>
- V.E. Tarasov, Exact discretization by Fourier transforms. *Commun. Nonlinear Sci. Numer. Simul.* **37**, 31–61 (2016f)
- V.E. Tarasov, United lattice fractional integro-differentiation. *Frac. Calc. Appl. Anal.* **19**(3), 625–664 (2016g). <https://doi.org/10.1515/fca-2016-0034>
- V.E. Tarasov, Exact discretization of Schrodinger equation. *Phys. Lett. A*. **380**(1–2), 68–75 (2016h)
- V.E. Tarasov, What discrete model corresponds exactly to gradient elasticity equation?. *J. Mech. Mater. Struct.* **11**(4), 329–343 (2016i). <https://doi.org/10.2140/jomms.2016.11.329>
- V.E. Tarasov, Exact solution of T-difference radial Schrodinger equation. *Int. J. Appl. Comput. Math.* (2017a). <https://doi.org/10.1007/s40819-016-0270-8>
- V.E. Tarasov, Exact discretization of fractional Laplacian. *Comput. Math. Appl.* **73**(5), 855–863 (2017b). <https://doi.org/10.1016/j.camwa.2017.01.012>
- V.E. Tarasov, J.J. Trujillo, Fractional power-law spatial dispersion in electrodynamics. *Ann. Phys.* **334**, 1–23 (2013). arXiv:1503.04349
- V.E. Tarasov, G.M. Zaslavsky, Fractional dynamics of coupled oscillators with long-range interaction. *Chaos*. **16**(2), 023110 (2006a). arXiv:nlin.PS/0512013
- V.E. Tarasov, G.M. Zaslavsky, Fractional dynamics of systems with long-range interaction. *Commun. Nonlinear Sci. Numer. Simul.* **11**(8), 885–898 (2006b). arXiv:1107.5436



Fractional Differential Calculus and Continuum Mechanics

24

K. A. Lazopoulos and A. K. Lazopoulos

Contents

Introduction	854
Basic Properties of Fractional Calculus	858
The Geometry of Fractional Differential	860
Differentiation	863
Linear Combinations	863
Linear Maps	864
Bilinear Maps	864
Cartesian Products	864
Compositions	864
The Fractional Arc Length	865
The Fractional Tangent Space	866
Fractional Curvature of Curves	867
The Fractional Radius of Curvature of a Curve	867
The Serret-Frenet Equations	868
Applications	870
The Fractional Geometry of a Parabola	870
The Tangent and Curvature Center of the Weierstrass Function	872
Bending of Fractional Beams	872
The Fractional Tangent Plane of a Surface	873
Fundamental Differential Forms on Fractional Differential Manifolds	875
The First Fractional Fundamental Form	875

The present essay is dedicated to Pepi Lazopoulou M.D., adorable wife and mother of the authors for her devotion in our family

K. A. Lazopoulos (✉)
National Technical University of Athens, Rafina, Greece
e-mail: kolazop@mail.ntua.gr

A. K. Lazopoulos
Mathematical Sciences Department, Hellenic Army Academy, Vari, Greece
e-mail: Orfeakos74@gmail.com

The Second Fractional Fundamental Form	876
The Fractional Normal Curvature	877
Fractional Vector Operators	878
Fractional Vector Field Theorems	879
Fractional Green's Formula	879
Fractional Stoke's Formula	880
Fractional Gauss' Formula	881
Fractional Deformation Geometry	881
Polar Decomposition of the Deformation Gradient	883
Deformation of Volume and Surface	884
Examples of Deformations	885
Homogeneous Deformations	885
The Nonhomogeneous Deformations	887
The Infinitesimal Deformations	888
Fractional Stresses	888
The Balance Principles	889
Material Derivatives of Volume, Surface, and Line Integrals	889
The Balance of Mass	891
Balance of Linear Momentum Principle	892
Balance of Rotational Momentum Principle	892
Fractional Zener Viscoelastic Model	893
The Integer Viscoelastic Model	893
The Fractional Order Derivative Viscoelastic Models	894
Proposed Fractional Viscoelastic Zener Model	895
Comparison of the Three Viscoelastic Models	898
Conclusion: Further Research	900
References	902

Abstract

The present essay is an attempt to present a meaningful continuum mechanics formulation into the context of fractional calculus. The task is not easy, since people working on various fields using fractional calculus take for granted that a fractional physical problem is set up by simple substitution of the conventional derivatives to any kind of the plethora of fractional derivatives. However, that procedure is meaningless, although popular, since laws in science are derived through differentials and not through derivatives. One source of that mistake is that the fractional derivative of a variable with respect to itself is different from one. The other source of the same mistake is that the well-known derivatives are not able to form differentials. This leads to erroneous and meaningless quantities like fractional velocity and fractional strain. In reality those quantities, that nobody understands what physically represent, alter the dimensions of the physical quantities. In fact the dimension of the fractional velocity is L/T^α , contrary to the conventional L/T . Likewise, the dimension of the fractional strain is $L^{-\alpha}$, contrary to the conventional L^0 . That fact cannot be justified. Imagine that even in relativity theory, where everything is changed, like time, lengths, velocities, momentums, etc., the dimensions remain constant. Fractional calculus is allowed up to now to change the dimensions and to accept derivatives

that are not able to form differentials, according to differential topology laws. Those handicaps have been pointed out in two recent conferences dedicated to fractional calculus by the authors, (K.A. Lazopoulos, in *Fractional Vector Calculus and Fractional Continuum Mechanics, Conference "Mechanics though Mathematical Modelling", celebrating the 70th birthday of Prof. T. Atanackovic*, Novi Sad, 6–11 Sept, Abstract, p. 40, 2015; K.A. Lazopoulos, A.K. Lazopoulos, Fractional vector calculus and fractional continuum mechanics. *Prog. Fract. Diff. Appl.* **2**(1), 67–86, 2016a) and were accepted by the fractional calculus community. The authors in their lectures (K.A. Lazopoulos, in *Fractional Vector Calculus and Fractional Continuum Mechanics, Conference "Mechanics though Mathematical Modelling", celebrating the 70th birthday of Prof. T. Atanackovic*, Novi Sad, 6–11 Sept, Abstract, p. 40, 2015; K.A. Lazopoulos, in *Fractional Differential Geometry of Curves and Surfaces, International Conference on Fractional Differentiation and Its Applications (ICFDA 2016)*, Novi Sad, 2016b; A.K. Lazopoulos, On *Fractional Peridynamic Deformations, International Conference on Fractional Differentiation and Its Applications, Proceedings ICFDA 2016*, Novi Sad, 2016c) and in the two recently published papers concerning fractional differential geometry of curves and surfaces (K.A. Lazopoulos, A.K. Lazopoulos, On the fractional differential geometry of curves and surfaces. *Prog. Fract. Diff. Appl.*, No **2**(3), 169–186, 2016b) and fractional continuum mechanics (K.A. Lazopoulos, A.K. Lazopoulos, Fractional vector calculus and fractional continuum mechanics. *Prog. Fract. Diff. Appl.* **2**(1), 67–86, 2016a) added in the plethora of fractional derivatives one more, that called Leibniz L-fractional derivative. That derivative is able to yield differential and formulate fractional differential geometry. Using that derivative the dimensions of the various quantities remain constant and are equal to the dimensions of the conventional quantities. Since the establishment of fractional differential geometry is necessary for dealing with continuum mechanics, fractional differential geometry of curves and surfaces with the fractional field theory will be discussed first. Then the quantities and principles concerning fractional continuum mechanics will be derived. Finally, fractional viscoelasticity Zener model will be presented as application of the proposed theory, since it is of first priority for the fractional calculus people. Hence the present essay will be divided into two major chapters, the chapter of fractional differential geometry, and the chapter of the fractional continuum mechanics. It is pointed out that the well-known historical events concerning the evolution of the fractional calculus will be circumvented, since the goal of the authors is the presentation of the fractional analysis with derivatives able to form differentials, formulating not only fractional differential geometry but also establishing the fractional continuum mechanics principles. For instance, following the concepts of fractional differential and Leibniz's L-fractional derivatives, proposed by the author (K.A. Lazopoulos, A.K. Lazopoulos, Fractional vector calculus and fractional continuum mechanics. *Prog. Fract. Diff. Appl.* **2**(1), 67–86, 2016a), the L-fractional chain rule is introduced. Furthermore, the theory of curves and surfaces is revisited, into the context of fractional calculus. The fractional tangents, normals, curvature vectors, and radii of curvature of curves

are defined. Further, the Serret-Frenet equations are revisited, into the context of fractional calculus. The proposed theory is implemented into a parabola and the curve configured by the Weierstrass function as well. The fractional bending problem of an inhomogeneous beam is also presented, as implementation of the proposed theory. In addition, the theory is extended on manifolds, defining the fractional first differential (tangent) spaces, along with the revisiting first and second fundamental forms for the surfaces. Yet, revisited operators like fractional gradient, divergence, and rotation are introduced, outlining revision of the vector field theorems. Finally, the viscoelastic mechanical Zener system is modelled with the help of Leibniz fractional derivative. The compliance and relaxation behavior of the viscoelastic systems is revisited and comparison with the conventional systems and the existing fractional viscoelastic systems are presented.

Keywords

Fractional Derivative · Fractional Differential · Fractional Stress · Fractional Strain · Fractional Principles in Mechanics · Fractional Continuum Mechanics

Introduction

Fractional Calculus, originated by Leibnitz (1849), Liouville (1832), and Riemann (1876) has recently applied to modern advances in physics and engineering. Fractional derivative models account for long-range (nonlocal) dependence of phenomena, resulting in better description of their behavior. Various material models, based upon fractional time derivatives, have been presented, describing their viscoelastic interaction (Atanackovic 2002; Mainardi 2010). Lazopoulos (2006) has proposed an elastic uniaxial model, based upon fractional derivatives for lifting Noll's axiom of local-action (Carpinteri et al. 2011) have also proposed a fractional approach to nonlocal mechanics. Applications in various physical areas may also be found in various books (Kilbas et al. 2006; Samko et al. 1993; Poldubny 1999; Oldham and Spanier 1974).

Since the need for Fractional Differential Geometry has extensively been discussed in various places, researchers have presented different aspects, concerning fractional geometry of manifolds (Tarasov 2010; Calcani 2012) with applications in fields of mechanics, quantum mechanics, relativity, finance, probabilities, etc. Nevertheless, researchers are raising doubtfulness about the existence of fractional differential geometry and their argument is not easily rejected.

Basically, the classical differential $df(x) = f'(x)dx$ has been substituted by the fractional one introduced by Adda (2001, 1998) in the form:

$$d^a f = g(x)(dx)^a$$

Nevertheless that definition of the differential is valid in the case of positive increments dx , whereas in the case of negative increments, the differential $d^a f(x)$

may be complex. That is exactly the reason why many researchers reasonably reject the existence of fractional differential geometry. However the variable x accepts its own fractional differential:

$$d^\alpha x = \sigma(x)(dx)^\alpha$$

with $\sigma(x) \neq 1$, differently of the conventional case when $a = 1$, where $\sigma(x)$ is always one. Relating both equations, it appears that:

$$d^a f = \frac{g(x)}{\sigma(x)} d^a x$$

In this case $d^\alpha x$ is always a real quantity accepting positive or negative incremental real values alike. On these bases, the development of fractional differential geometry may be established.

Further, fractal functions exhibiting self-similarity are nondifferentiable functions, but they exhibit fractional differentiability of order $0 < \alpha < 1$ (see Yao et al. 2005; Carpinteri et al. 2009; Goldmankhaneh et al. 2013; Liang and Su 2007). Goldmankhaneh et al. (2013) introduced the generalized fractional Riemann-Liouville and Caputo-like derivatives for functions defined on fractal sets.

Fractional Calculus in mechanics has been suggested by many researchers, Tarasov (2010, 2008), Drapaca and Sivaloganathan (2012), Sumelka (2014), and Lazopoulos and Lazopoulos (2016a), in problems of continuum mechanics with microstructure where nonlocal elasticity is necessary. Fractional continuum mechanics has been applied to various problems in hydrodynamics (Tarasov 2010; Balankin and Elizarrataz 2012). Recently fractional calculus has been introduced by the author (Lazopoulos 2016a) for the description of peridynamic theory (Silling 2003, 2010). Yet, fractional calculus has been considered as the best frame for describing viscoelastic problems (Atanackovic 2002; Beyler and Kempfle 1995). In addition fractional differential geometry affects rigid body dynamics, in holonomic and nonholonomic systems (Riewe 1996, 1997; Baleanu et al. 2013). Recent applications in quantum mechanics, physics, and relativity demand differential geometry revisited by fractional calculus (Golmankhaneh Ali et al. 2015; Baleanu et al. 2009).

In the present work, the fractional differential established in Lazopoulos and Lazopoulos (2016a) will be recalled along with the introduced Leibniz's L-fractional derivatives. Those differentials are always real and proper for establishing the fractional differential geometry. Correcting the picture of fractional differential of a function, the fractional tangent space of a manifold was defined, introducing also Leibniz's L-fractional derivative that is the only one having physical meaning. Moreover, the present work reviews description of fractional geometry of curves, describing their tangent (first differential) spaces, their normals, the curvature vectors, and the corresponding radii of curvature. In addition, the Serret-Frenet equations will be revisited into the fractional calculus context. The theory is implemented to a parabola, to the Weierstrass function and the beam bending (Lazopoulos et al. 2015), considered as applications of the curves' theory to the solid

mechanics. Yet, the theory is extended on manifolds, just to describe the fractional differential geometry of surfaces. Finally outline of fractional vector field theory is included, along with the revisited fractional vector field theorems.

The mechanics researchers have been motivated by the mechanical behavior of disordered (nonhomogeneous) materials with microstructure. Porous materials (Vardoulakis et al. 1998, Ma et al. 2002), colloidal aggregates (Wyss et al. 2005), ceramics, etc., are materials with microstructure that exert strong influence in their deformation. Major factors in determining the material deformation are microcracks, voids, material phases, etc. The nonhomogeneity of the heterogeneous materials has been tackled by various homogenization theories (Bakhalov and Panasenko 1989). Nevertheless, these materials require the lifting of the basic local action axiom of continuum mechanics (Truesdell 1977; Truesdell and Noll 1965). As defined by Noll (1958, 1959) simple materials satisfy the three fundamental axioms:

- (a) The principle of determinism
- (b) The principle of local action
- (c) The principle of material frame-indifference.

Truesdell and Noll (1965) points out in his classic continuum mechanics book: “The motion of body-points at a finite distance from a point x in some shape may be disregarded in calculating the stress at x .” Material microstructure, inhomogeneities, microcracks, etc., are some of the various important factors that affect the material deformation with nonlocal action. These factors are not considered in the simple materials formulation.

Various theories have been proposed just to introduce a long distance action in the deformation of the materials. One direction considers Taylor’s expansion of the strain tensor in the neighborhood of a point, taking in consideration one or two most important terms. Hence gradient strain theories have been appeared in nonlinear form (Toupin 1965), and in linear deformation (Mindlin 1965). Eringen (2002) has also proposed a theory dealing with micropolar elasticity. Mindlin introduced a simpler version of linear gradient theories and an even simpler model has been presented by Aifantis (1999) with his GRADELA model. In these theories, the authors introduced intrinsic material lengths that accompany the higher order derivatives of the strain. Many problems have been solved employing those theories concerning size effects, lifting of various singularities, porous materials (Aifantis 1999, 2003, 2011; Askes and Aifantis 2011), mechanics of microbeams, microplates, and microsheets (Lazopoulos 2004; Lazopoulos et al. 2010).

However various nonlocal elastic theories have been introduced, that are more reliable in taking care of nonsmooth deformations, since integrals are friendlier than derivatives to take care of various nonsmooth phenomena.

Lazopoulos (2006) introduced fractional derivatives of the strain in the strain energy density function in an attempt to introduce nonlocality in the elastic response of materials. Fractional calculus was used by many researchers, not only in the field of Mechanics but mainly in Physics and especially in Quantum Mechanics, to develop the idea of introducing nonlocality. In fact, the history of fractional calculus

is dated since seventeenth century. Particle physics, electromagnetics, mechanics of materials, hydrodynamics, fluid flow, rheology, viscoelasticity, optics, electrochemistry and corrosion, and chemical physics are some fields where fractional calculus has been introduced.

Fractional calculus in material deformations has been adopted in solving various types of problems. First we may consider the deformation problems with nonsmooth strain field. Second heterogeneous material deformations may also be studied. Furthermore, time fractional derivative is proved to be more suitable in viscoelastic deformations, since viscoelastic deformations with retarded memory materials may also be discussed. The nonlocal strain effects of deformation problems are concerned by the last type of those problems. There are many studies considering fractional elasticity theory, introducing fractional strain (Drapaca and Sivaloganathan 2012; Carpinteri et al. 2001, 2011; Di Paola et al. 2009; Atanackovic et al. 2008; Agrawal 2008). Recently Sumelka (2014) has presented applications of Fractional Calculus in the nonlocal elastic deformation of Kirchhoff-Love plates and in the rate-independent plasticity. Nevertheless a different definition of fractional strain is yielded in the present work. Jumarie (2012) has proposed modified Riemann-Liouville derivative of fractional order with an approach to differential geometry of fractional order. In addition Meerschaert et al. (2006) have presented fractional order vector calculus for fractional advection-dispersion. Recent applications of fractional calculus have been appeared in peridynamic theory (Silling 2000; Silling et al. 2003; Evangelatos 2011; Evangelatos and Spanos 2012). Tarasov (2010) has also presented a book including fractional mathematics and its applications to various physics areas. In addition Tarasov (2008) has presented a fractional vector fields theory combining fractals (Feder 1988), and fractional calculus.

Lazopoulos and Lazopoulos (2016a, b) have clarified the geometry of the fractional differential resulting in fractional tangent spaces of the manifolds quite different from the conventional ones. Hence the fractional differential geometry has been established, indispensable for the development of fractional mechanics. It is evident that the definition of the stress and the strain is greatly affected by the tangent spaces. Hence the fractional stress tensors and the fractional strain tensors are quite different from the conventional ones. The linear strain tensors are also revisited. Those basic concepts are important for establishing fractional continuum mechanics.

In the present work, fractional vector calculus is revisited, since the fractional differential of a function is not linearly dependent upon the conventional differential of the variables. Furthermore, the fractional derivative of a variable with respect to itself is different from one. The fractional vector calculus is revisited along with the basic field theorems of Green, Stokes, and Gauss. Applications of the fractional vector calculus to continuum mechanics are presented. The revision in the right and left Cauchy-Green deformation tensors and Green (Lagrange) and Euler-Almansi strain tensors are exhibited. The change of volume and the surface due to deformation (change of configuration) of a deformable body is also discussed. Further the revisited fractional continuum mechanics principles yielding the fractional continuity and motion equations are also derived.

Moreover, linearization of the strain tensors is performed. The change of fractional volume and fractional surface due to deformation (change of configuration) of a deformable body is also discussed. In addition, the revisited fractional continuum mechanics principles, such as mass conservation (fractional continuity equation) and motion equations (conservation of fractional linear and rotational momentum), are also discussed.

In addition, Bagley et al. (1986) introduced Fractional Calculus in viscoelasticity, and Atanackovic et al. (2002, 2002a) pursue the idea in many applications and in fractional variational problems (Atanackovic et al. 2008) with fractional derivatives. Mainardi (2010) has also discussed the application of fractional calculus in linear viscoelasticity. In Sabatier et al. (2007), there exists a section concerning viscoelastic disordered media. As an application to the present theory, the viscoelastic behavior of Zener model will be revisited using Leibniz fractional time derivatives. Comparison of the proposed model to the existing fractional ones will be discussed. Further the behavior of the proposed model concerning its compliance and relaxation is discussed and compared to the existing fractional ones and the conventional as well.

It is pointed out that the present essay will be divided into two major chapters, the chapter of fractional differential geometry and the chapter of the fractional continuum mechanics. The viscoelasticity Zener model will be discussed into the context of the proposed theory.

Basic Properties of Fractional Calculus

Fractional Calculus has recently become a branch of pure mathematics, with many applications in Physics and Engineering, (Tarasov 2008, 2010). Many definitions of fractional derivatives exist. In fact, fractional calculus originated by Leibniz, looking for the possibility of defining the derivative $\frac{d^n g}{dx^n}$ when $n = \frac{1}{2}$. The various types of the fractional derivatives exhibit some advantages over the others. Nevertheless they are almost all nonlocal, contrary to the conventional ones.

The detailed properties of fractional derivatives may be found in Kilbas et al. (2006), Podlubny (1999), and Samko et al. (1993). Starting from Cauchy formula for the n -fold integral of a primitive function $f(x)$

$$I^n f(x) = \int_a^x f(s) (ds)^n = \int_a^x dx_n \int_a^{x_n} dx_{n-1} \int_a^{x_{n-1}} dx_{n-2} \dots \int_a^{x_2} f(x_1) dx_1 \quad (1)$$

expressed by:

$${}_a I_x^n f(x) = \frac{1}{(n-1)!} \int_a^x (x-s)^{n-1} f(s) ds, \quad x > 0, n \in N \quad (2)$$

and

$${}_x I_b^n f(x) = \frac{1}{(n-1)!} \int_x^b (s-x)^{n-1} f(s) ds, \quad x > 0, n \in \mathbb{N} \tag{3}$$

the left and right fractional integral of f are defined as:

$${}_a I_x^\alpha f(x) = \frac{1}{\Gamma(\alpha)} \int_a^x \frac{f(s)}{(x-s)^{1-\alpha}} ds \tag{4}$$

$${}_x I_b^\alpha f(x) = \frac{1}{\Gamma(\alpha)} \int_x^b \frac{f(s)}{(s-x)^{1-\alpha}} ds \tag{5}$$

In Eqs. 4 and 5 we assume that α is the order of fractional integrals with $0 < \alpha \leq 1$, considering $\Gamma(x) = (x-1)!$ with $\Gamma(\alpha)$ Euler’s Gamma function.

Thus the left and right Riemann-Liouville (R-L) derivatives are defined by:

$${}_a D_x^\alpha f(x) = \frac{d}{dx} ({}_a I_x^{1-\alpha} f(x)) \tag{6}$$

and

$${}_x D_b^\alpha f(x) = -\frac{d}{dx} ({}_b I_x^{1-\alpha} f(x)) \tag{7}$$

Pointing out that the R-L derivatives of a constant c are nonzero, Caputo’s derivative has been introduced, yielding zero for any constant. Thus, it is considered as more suitable in the description of physical systems.

In fact Caputo’s derivative is defined by:

$${}_a^c D_x^\alpha f(x) = \frac{1}{\Gamma(1-\alpha)} \int_a^x \frac{f'(s)}{(x-s)^\alpha} ds \tag{8}$$

$$\text{and } {}_x^c D_b^\alpha f(x) = -\frac{1}{\Gamma(1-\alpha)} \int_x^b \frac{f'(s)}{(s-x)^\alpha} ds \tag{9}$$

Evaluating Caputo’s derivatives for functions of the type:

$f(x) = (x-a)^n$ or $f(x) = (b-x)^n$ we get:

$${}_a^c D_x^\alpha (x-a)^\nu = \frac{\Gamma(\nu+1)}{\Gamma(-\alpha+\nu+1)} (x-a)^{\nu-\alpha}, \tag{10}$$

and for the corresponding right Caputo’s derivative:

$${}_x D_b^{\alpha} (b - x)^{\nu} = \frac{\Gamma(\nu + 1)}{\Gamma(-\alpha + \nu + 1)} (b - x)^{\nu - \alpha}$$

Likewise, Caputo’s derivatives are zero for constant functions:

$$f(x) = c. \tag{11}$$

Finally, Jumarie’s derivatives are defined by,

$${}_a D_x^{\alpha} f(x) = \frac{1}{\Gamma(1 - \alpha)} \frac{d}{dx} \int_a^x \frac{f(s) - f(a)}{(x - s)^{\alpha}} ds$$

and

$${}_x D_b^{\alpha} f(x) = -\frac{1}{\Gamma(1 - \alpha)} \frac{d}{dx} \int_x^b \frac{f(s) - f(b)}{(s - x)^{\alpha}} ds$$

Although those derivatives are accompanied by some derivation rules that are not valid, the derivatives themselves are valid and according to our opinion are better than Caputo’s, since they accept functions less smooth than the ones for Caputo. Also Jumarie’s derivative is zero for constant functions, basic property, advantage, of Caputo derivative. Nevertheless, Caputo’s derivative will be employed in the present work, having in mind that Jumarie’s derivative may serve better our purpose.

The Geometry of Fractional Differential

It is reminded, the n -fold integral of the primitive function $f(x)$, Eq. 1, is

$$I^n f(x) = \int_a^x f(s) (ds)^n \tag{12}$$

which is real for any positive or negative increment ds . Passing to the fractional integral

$$I^{\alpha} (f(x)) = \int_a^x f(s) (ds)^{\alpha} \tag{13}$$

the integer n is simply substituted by the fractional number α . Nevertheless, that substitution is not at all straightforward. The major difference between passing from

Eqs. 11 to 12 is that although $(ds)^n$ is real for negative values of ds , $(ds)^\alpha$ is complex. Therefore, the fractional integral, Eq. 13, is not compact for any increment ds . Hence the integral of Eq. 13 is misleading. In other words, the differential, necessary for the existence of the fractional integral, Eq. 13, is wrong. Hence, a new fractional differential, real and valid for positive and negative values of the increment ds , should be established.

It is reminded that the α -Fractional differential of a function $f(x)$ is defined by, Adda (1998):

$$d^\alpha f(x) = {}^c D_x^\alpha f(x)(dx)^\alpha \tag{14}$$

It is evident that the fractional differential, defined by Eq. 14, is valid for positive incremental dx , whereas for negative ones, that differential might be complex. Hence considering for the moment that the increment dx is positive, and recalling that ${}^c D_x^\alpha x \neq 1$, the α -fractional differential of the variable x is:

$$d^\alpha x = {}^c D_x^\alpha x(dx)^\alpha \tag{15}$$

Hence

$$d^\alpha f(x) = \frac{{}^c D_x^\alpha f(x)}{{}^c D_x^\alpha x} d^\alpha x \tag{16}$$

It is evident that $d^\alpha f(x)$ is a nonlinear function of dx , although it is a linear function of $d^\alpha x$. That fact suggests the consideration of the fractional tangent space that we propose. Now the definition of fractional differential, Eq. 16, is imposed either for positive or negative variable differentials $d^\alpha x$. In addition the proposed L-fractional (in honor of Leibniz) derivative ${}^L D_x^\alpha f(x)$ is defined by,

$$d^\alpha f(x) = {}^L D_x^\alpha f(x) d^\alpha x \tag{17}$$

with the Leibniz L-fractional derivative,

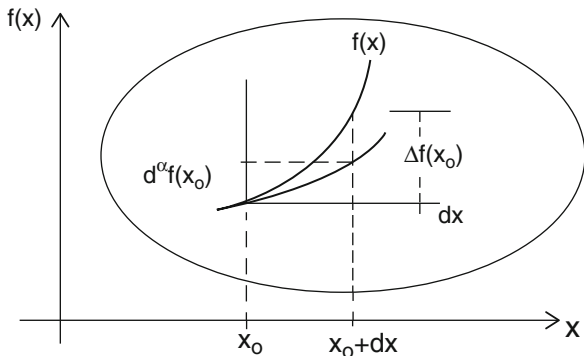
$${}^L D_x^\alpha f(x) = \frac{{}^c D_x^\alpha f(x)}{{}^c D_x^\alpha x} \tag{18}$$

Hence only Leibniz’s derivative has any geometrical of physical meaning.

In addition, Eq. 3, is deceiving and the correct form of Eq. 3 should be substituted by,

$$f(x) - f(a) = {}^L I_x^\alpha \left({}^L D_x^\alpha f(x) \right) = \frac{1}{\Gamma(\alpha) \Gamma(2-\alpha)} \int_a^x \frac{(s-a)^{1-\alpha}}{(x-s)^{1-\alpha}} {}^L D_x^\alpha f(s) ds \tag{19}$$

Fig. 1 The nonlinear differential of $f(x)$



It should be pointed out that the correct forms are defined for the fractional differential by Eq. 17, the Leibniz derivative, Eq. 18 and the fractional integral by Eq. 19. All the other forms are misleading.

Configuring the fractional differential, along with the first fractional differential space (fractional tangent space), the function $y = f(x)$ has been drawn in Fig. 1, with the corresponding first differential space at a point x , according to Adda’s definition, Eq. 14.

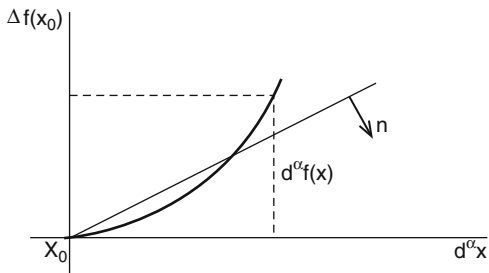
The tangent space, according to Adda’s (1998) definition, Eq. 14, is configured by the nonlinear curve $d^\alpha f(x)$ versus dx . Nevertheless, there are some questions concerning the correct picture of the configuration, (Fig. 1), concerning the fractional differential presented by Adda (1998). Indeed,

- (a) The tangent space should be linear. There is not conceivable reason for the nonlinear tangent spaces.
- (b) The differential should be configured for positive and negative increments dx . However, the tangent spaces, in the present case, do not exist for negative increments dx .
- (c) The axis $d^\alpha f(x)$, in Fig. 1, presents the fractional differential of the function $f(x)$, however dx denotes the conventional differential of the variable x . It is evident that both axes along x and $f(x)$ should correspond to differentials of the same order.

Therefore, the tangent space (first differential space), should be configured in the coordinate system with axes $(d^\alpha x, d^\alpha f(x))$. Hence, the fractional differential, defined by Eq. 17, is configured in the plane $(d^\alpha x, d^\alpha f(x))$ by a line, as it is shown in Fig. 2.

It is evident that the differential space is not tangent (in the conventional sense) to the function at x_0 , but intersects the figure $y = f(x)$ at least at one point x_0 . This space, we introduce, is the tangent space. Likewise, the normal is perpendicular to the line of the fractional tangent. Hence we are able to establish fractional differential geometry of curves and surfaces with the fractional field theory. Consequently when $\alpha = 1$, the tangent spaces, we propose, coincide with

Fig. 2 The virtual tangent space of the $f(x)$ at the point $x = x_0$



the conventional tangent spaces. As a last comment concerning the proposed L-fractional derivative, the physical dimensions of the various quantities remain unaltered from the conventional to any order Fractional Calculus.

Differentiation

The following chapter is a summary from the paragraphs 2.4–2.6 of the book, Differential topology with a view to applications (Chillingworth 1976).

Let E and F be two normed linear spaces with respective norms $\|\cdot\|_E$ and $\|\cdot\|_F$, suppose $f: E \rightarrow F$ is a given continuous (not necessarily linear) map, and let x be a particular point of E .

Definition

The map f is differentiable if there exists a linear map $L: E \rightarrow F$ which approximates f at x in the sense that for all h in E we have:

$$f(x + h) - f(x) = L(h) + \|h\|_E \eta(h) \tag{20}$$

where $h(x)$ is an element of F with,

$$\|\eta(h)\|_F \rightarrow 0 \text{ as } \|h\|_E \rightarrow 0 \tag{21}$$

In this definition, E and F are two normed linear spaces with respective norms $\|\cdot\|_E$ and $\|\cdot\|_F$, and $f: E \rightarrow F$ is a given continuous (not necessarily linear) map.

Furthermore, let x be a particular point of E and U an open subset of E . The properties and uses of the derivative are:

Linear Combinations

If $f, g: U \rightarrow F$ are differentiable then so is the map $\alpha f + \beta g$ for any constants α, β and $D(\alpha f + \beta g) = \alpha Df + \beta Dg$ as maps $U \rightarrow L(E, F)$. At this point we must point out that constant maps themselves have derivative zero.

Linear Maps

If f is the restriction to U of a continuous linear map: $E \rightarrow F$, then $Df(x) = L$ for every $x \in U$, i.e. L is already its own linear approximation.

In the special case $E = F = R$ we have $f(x) = \zeta x$ for some number ζ , and $Df(x) = \zeta$ regarded as the linear map $R \rightarrow R: x \rightarrow \zeta x$.

Bilinear Maps

If $E = E_1 \times E_2$ and f is the restriction to U of a continuous bilinear map $B: E_1 \times E_2 \rightarrow F$ (i.e. B is linear in each factor separately) then $Df(x)h = B(x_1, h_2) + B(h_1, x_2)$ where $x = (x_1, x_2)$, $h = (h_1, h_2)$ with x_1, h_1 in E_1 and x_2, h_2 in E_2 .

Cartesian Products

If $f_1: U_1 \rightarrow F_1$ and $f_2: U_2 \rightarrow F_2$ are differentiable then so is $f = f_1 \times f_2: U_1 \times U_2 \rightarrow F_1 \times F_2$ where $(f_1 \times f_2)(x_1, x_2)$ means $(f_1(x_1), f_2(x_2))$, and we have:

$$Df(x_1, x_2)(h_1 \times h_2) = (Df_1(x_1)h_1, Df_2(x_2)h_2) \quad (22)$$

That is, derivatives operate coordinate-wise.

Compositions

Chain Rule

Suppose E, F , and G are three normed linear spaces, and U, V are open sets in E, F , respectively. Let $f: U \rightarrow F$ and $g: V \rightarrow G$ be continuous maps, and suppose the image of f lies in V so that the composition $g \circ f: U \rightarrow G$ exists.

If f is differentiable at x and g is differentiable at $f(x)$ then $g \circ f$ is differentiable

$$D(g \circ f)(x) = Dg(f(x)) \bullet Df(x) \quad (23)$$

In other words, the derivative of a composition is the composition of the derivatives.

If $E = R^n, F = R^m, G = R^p$, and $y = f(x), z = z(y)$ the chain rule states that:

$$\frac{\partial (z_1, z_2, \dots, z_p)}{\partial (x_1, x_2, \dots, x_n)} = \frac{\partial (z_1, z_2, \dots, z_p)}{\partial (y_1, y_2, \dots, y_m)} \cdot \frac{\partial (y_1, y_2, \dots, y_m)}{\partial (x_1, x_2, \dots, x_n)} \quad (24)$$

Product Rule

For the same maps f, g holds:

$$D(f \cdot g) = f \cdot D(g) + g \cdot D(f) \tag{25}$$

Let's assume derivative of a product $f(t) = A(t) \cdot b(t)$ where $A(t)$ is an $m \times n$ matrix and $b(t)$ is an n vector. Then:

$$Df(t) = A(t) \cdot Db(t) + DA(t) \cdot b(t) \tag{26}$$

Now the question arises whether L-fractional derivative satisfies all the conditions, Eqs. 22, 23, 24, 25, and 26 required by the demand of differentiation according to differential topology rules. The answer is no. There are some rules that the L-fractional derivative satisfies, like the linearity condition, Eq. 17 that cannot be satisfied by the common fractional derivatives like Caputo etc., since the nonlinear Eq. 14 holds for them. However, there are other conditions, such as the chain rule that are not valid. Now we have to make a choice. Either to define differential that is necessary for establishing fractional differential geometry that is necessary for dealing with problems in Physics, or to forget fractional calculus. On that dilemma we make the choice to *impose* the necessary rules, like fractional chain rule, just for forming fractional differential geometry. In that case the physical problem leaves its print or trace on mathematics, that in conventional calculus is not necessary, since all differential rules are satisfied by themselves, in the conventional case.

The Fractional Arc Length

Let $y = f(x)$ be a function, which may be non-differentiable but has a fractional derivative of order α , $0 < \alpha < 1$. The fractional differential of $y = f(x)$ in the differential space is defined by:

$$d^a y = \frac{{}_a D_x^a f(x)}{{}_a D_x^a x} d^a x = {}_a^L D_x^a f(x) d^a x \tag{27}$$

Therefore the arc length is expressed by:

$$s_1(x, a) = {}_a I_x^a \left[(d^a y)^2 + (d^a x)^2 \right]^{1/2} = {}_a I_x^a \left[\left(\frac{{}_a D_x^a f(x)}{{}_a D_x^a x} \right)^2 + 1 \right]^{\frac{1}{2}} d^a x \tag{28}$$

Furthermore, for parametric curves of the type:

$$y = f(t), x = g(t) \tag{29}$$

The fractional α -differentials are defined by:

$$\begin{aligned} d^a x &= \frac{{}_a D_t^\alpha g(t)}{{}_a D_t^\alpha t} d^a t \\ d^a y &= \frac{{}_a D_t^\alpha f(t)}{{}_a D_t^\alpha t} d^a t \end{aligned} \tag{30}$$

and the fractional differential of the arc length is expressed by:

$$d^a s = \sqrt{(d^a y)^2 + (d^a x)^2} = \left[\left(\frac{{}_a D_t^\alpha f(t)}{{}_a D_t^\alpha t} \right)^2 + \left(\frac{{}_a D_t^\alpha g(t)}{{}_a D_t^\alpha t} \right)^2 \right]^{\frac{1}{2}} d^a t \tag{31}$$

and

$$\begin{aligned} s &= {}^L I_x^a d^a s = {}^L I_t^a \left[\left(\frac{{}_a D_t^\alpha f(t)}{{}_a D_t^\alpha t} \right)^2 + \left(\frac{{}_a D_t^\alpha g(t)}{{}_a D_t^\alpha t} \right)^2 \right]^{\frac{1}{2}} d^a t \\ &= {}^L I_t^a \left[\left({}^L D_t^\alpha f(t) \right)^2 + \left({}^L D_t^\alpha g(t) \right)^2 \right]^{\frac{1}{2}} d^a t \end{aligned} \tag{32}$$

The Fractional Tangent Space

Let $\mathbf{r} = \mathbf{r}(s)$ be a natural representation of a curve C , where s is the α -fractional length of the curve. Since the velocity of a moving material point on the curve $r(s)$ defines the tangent space, the fractional tangent space of the curve $\mathbf{r} = \mathbf{r}(s)$ is defined by the first derivative:

$$\mathbf{r}_1 = \frac{d^a \mathbf{r}}{d^a s} = \frac{{}_a D_s^a \mathbf{r}}{{}_a D_s^a s} = {}^L D_s^a \mathbf{r} \tag{33}$$

Recalling

$$d^a |\mathbf{r}| = d^a s \tag{34}$$

the length $|\mathbf{r}_1|$ of the fractional tangent vector is unity.

The tangent space line of the curve $\mathbf{r} = \mathbf{r}(s)$ at the point $\mathbf{r}_0 = \mathbf{r}(s_0)$ is defined by:

$$\mathbf{r} = \mathbf{r}_0 + k \mathbf{t}_0 \quad 0 < k < \infty \tag{35}$$

where $\mathbf{t}_0 = \mathbf{t}(s_0)$ is the unit tangent vector at \mathbf{r} .

The plane through \mathbf{r}_0 , orthogonal to the tangent line at \mathbf{r}_0 , is called the normal plane to the curve C at s_0 . The points \mathbf{y} of that orthogonal plane are defined by:

$$(\mathbf{y} - \mathbf{r}_0) \cdot \mathbf{t}(s_0) = (\mathbf{y} - \mathbf{r}_0) \cdot \mathbf{r}_1(s_0) = 0 \tag{36}$$

Fractional Curvature of Curves

Considering the fractional tangent vector:

$$\mathbf{t} = \mathbf{r}_1(s) = \frac{{}_a D_s^a \mathbf{r}}{{}_a D_s^a s} = {}^L D_s^a \mathbf{r} \quad (37)$$

its fractional derivative may be considered:

$$\mathbf{r}_2(s) = \frac{d^a \mathbf{t}}{d^a s} = \frac{{}_a D_s^a \mathbf{t}}{{}_a D_s^a s} = {}^L D_s^a \mathbf{t} = \mathbf{t}_1(s) \quad (38)$$

The vector $\mathbf{t}_1(s)$ is called the fractional curvature vector on C at the point $\mathbf{r}(s)$ and is denoted by $\kappa = \kappa(s) = \mathbf{t}_1(s)$.

Since \mathbf{t} is a unit vector

$$\mathbf{t} \cdot \mathbf{t} = 1 \quad (39)$$

Restricted to fractional derivatives that yield zero for a constant function, such as Caputo's derivatives, the curvature vector $\mathbf{t}_1(s)$ on C is orthogonal to \mathbf{t} and parallel to the normal plane. The magnitude of the fractional curvature vector:

$$\kappa = |\kappa(s)| \quad (40)$$

is called the fractional curvature of C at $\mathbf{r}(s)$. The reciprocal of the curvature κ is the fractional radius of curvature at $\mathbf{r}(s)$:

$$\rho = \frac{1}{\kappa} = \frac{1}{|\kappa(s)|} \quad (41)$$

The Fractional Radius of Curvature of a Curve

Following Porteous (1994), for the fractional curvature of a plane curve \mathbf{r} , we study at each point $\mathbf{r}(t)$ of the curve, how closely the curve approximates there to a parameterized circle. Now in the tangent or first differential space at a point $\mathbf{r}(t_0)$, the circle, with center \mathbf{c} and radius ρ , consists of all $\mathbf{r}(t)$ in the differential space such that:

$$(\mathbf{r} - \mathbf{c}) \cdot (\mathbf{r} - \mathbf{c}) = \rho^2 \quad (42)$$

Further Eq. 42 yields:

$$\mathbf{c} \cdot \mathbf{r} - \frac{1}{2} \mathbf{r} \cdot \mathbf{r} = \frac{1}{2} (\mathbf{c} \cdot \mathbf{c} - \rho^2) \quad (43)$$

with the right hand side been constant.

Therefore the derivation of the function

$$\mathbf{V}(\mathbf{c}) : t \rightarrow \mathbf{c} \cdot \mathbf{r}(t) - \frac{1}{2} \mathbf{r}(t) \cdot \mathbf{r}(t) \tag{44}$$

Hence:

$$V(\mathbf{c})_1 = (\mathbf{c} - \mathbf{r}(t)) \cdot \mathbf{r}_1(t) = 0 \tag{45}$$

$$V(\mathbf{c})_2 = (\mathbf{c} - \mathbf{r}(t)) \cdot \mathbf{r}_2(t) - \mathbf{r}_1(t) \cdot \mathbf{r}_1(t) = 0 \tag{46}$$

Suppose that \mathbf{r} is a parametric curve with $\mathbf{r}(t)$ in the virtual tangent space. Then $V(\mathbf{c})_1(t) = 0$ when the vector $\mathbf{c} - \mathbf{r}(t)$ in the tangent space is orthogonal to the tangent vector $\mathbf{r}_1(t)$. Indeed when the point \mathbf{c} , in the tangent space, lies on the normal to $\mathbf{r}_1(t)$ at t , the line through $\mathbf{r}(t)$ is orthogonal to the tangent line.

When $\mathbf{r}_2(t)$ is not linearly dependent upon $\mathbf{r}_1(t)$, there will be a unique point $\mathbf{c} \neq \mathbf{r}(t)$, on the normal line, such that also $V(\mathbf{c})_2(t) = 0$.

The Serret-Frenet Equations

Let \mathbf{r} be a curve with unit speed, where the fractional velocity vector, (Porteous 1994),

$$\mathbf{t}(s) = \mathbf{r}_1(s) = \frac{{}_a^c D_s^a \mathbf{r}(s)}{{}_a^c D_s^a s} = {}_a^L D_s^a \mathbf{r}(s) \tag{47}$$

is of unit length.

Let $\mathbf{r}(s)$ be such a curve. The vector

$$\mathbf{t}_1(s) = \mathbf{r}_2(s) = \frac{{}_a^c D_s^a \mathbf{r}_1(s)}{{}_a^c D_s^a s} = \frac{{}_a^c D_s^a}{{}_a^c D_s^a s} \left(\frac{{}_a^c D_s^a \mathbf{r}(s)}{{}_a^c D_s^a s} \right) = {}_a^L D_s^a \left({}_a^L D_s^a \mathbf{r}(s) \right) \tag{48}$$

is normal to the curve $\mathbf{r} = \mathbf{r}(s)$ since $\mathbf{t}(s) \cdot \mathbf{r}(s) = 1$ and

$$\mathbf{t}_1(s) \cdot \mathbf{t}(s) = 0 \tag{49}$$

since for Caputo's derivative ${}_a^L D_s^a c = 0$ for any constant c .

Consider $\mathbf{t}_1(s) \kappa(s) \mathbf{n}(s)$, where $\mathbf{n}(s)$ is the unit principal normal to \mathbf{r} at s , provided that $\kappa(s) \neq 0$ where $\kappa(s)$ is the curvature of \mathbf{r} at s .

Hence the equations for the focal line are defined by

$$\begin{aligned}(\mathbf{c} - \mathbf{r}(s)) \cdot \mathbf{r}_1(s) &= 0 \\(\mathbf{c} - \mathbf{r}(s)) \cdot \kappa(s)\mathbf{n}(s) &= 1\end{aligned}\tag{50}$$

Thus, the principal center of curvature \mathbf{c} at s is the point $\mathbf{r}(s) + \rho(s)\mathbf{n}(s)$, where $\rho(s) = \frac{1}{\kappa(s)}$. Furthermore, the principal normal vector $\mathbf{n}(s)$ orthogonal to the tangent line is pointing towards the focal line (locus of the curvature centers). Likewise, the (unit) binormal $\mathbf{b}(s)$ is defined to be the vector $\mathbf{t}(s) \times \mathbf{n}(s)$, the triad of unit vectors $\mathbf{t}(s)$, $\mathbf{n}(s)$, $\mathbf{b}(s)$ forming a right-handed orthonormal basis for the tangent vector space to the curvature $\mathbf{r}(s)$. Each of the derivative vectors $\mathbf{t}_1(s)$, $\mathbf{n}_1(s)$, $\mathbf{b}_1(s)$ linearly depends on $\mathbf{t}(s)$, $\mathbf{n}(s)$, $\mathbf{b}(s)$. Considering the equations: $\mathbf{t}_1 \mathbf{t} = 0$ and $\mathbf{t}_1 \mathbf{n} = 0$ with $\mathbf{t}_1 \mathbf{n} + \mathbf{n}_1 \cdot \mathbf{t} = 0$, we get the fractional Serret-Frenet equations:

$$\begin{aligned}\mathbf{t}_1 &= \kappa \mathbf{n} \\ \mathbf{n}_1 &= -\kappa \mathbf{t} + \tau \mathbf{b} \\ \mathbf{b}_1 &= -\tau \mathbf{n}\end{aligned}\tag{51}$$

The coefficient τ is defined to be the torsion of the curve \mathbf{r} . These equations are the fractional equations for the fractional Serret-Frenet system. Considering plane curves,

$$\mathbf{r}(x) = x\mathbf{i} + y(x)\mathbf{j}\tag{52}$$

Equations 45 and 46 defining the fractional centers of curvature $\mathbf{c} = c_1\mathbf{i} + c_2\mathbf{j}$ become,

$$\begin{aligned}(c_1 - x) + (c_2 - y(x)) {}^L D_x^\alpha y(x) &= 0 \\ (c_2 - y(x)) {}^L D_x^\alpha ({}^L D_x^\alpha y(x)) - \left(1 + {}^L D_x^\alpha y(x)^2\right) &= 0\end{aligned}\tag{53}$$

Since the fractional radius of curvature is defined by

$$\rho^\alpha = \rho_1^\alpha \mathbf{i} + \rho_2^\alpha \mathbf{j} = (c_1 - x)\mathbf{i} + (c_2 - y(x))\mathbf{j}\tag{54}$$

the components of the fractional curvature are given by

$$\begin{aligned}\rho_1^\alpha &= -\frac{1 + {}^L D_s^\alpha y(x)^2}{{}^L D_s^\alpha ({}^L D_s^\alpha y(x))} {}^L D_s^\alpha y(x) \\ \rho_2^\alpha &= \frac{1 + {}^L D_s^\alpha y(x)^2}{{}^L D_s^\alpha ({}^L D_s^\alpha y(x))}\end{aligned}\tag{55}$$

Further, for the case of $|y(x)| \ll 1$, that we consider in linear bending, Eq. 28 yields,

$$d^a s = d^a x + o(d^a x)^2 \tag{56}$$

with

$${}^L D_s^\alpha () = {}^L D_x^\alpha () \tag{57}$$

and

$$\rho^\alpha = |\mathbf{r}^\alpha| \approx \frac{1}{{}^L D_x^\alpha ({}^L D_x^\alpha y(x))} \tag{58}$$

Let us consider a fractional beam with its source point $(x,y,z) = (0,0,0)$. That means, the fractional u Caputo's derivatives of any function, concerning the beam, are defined by

$${}_0^c D_u^a f(u) = \frac{1}{\Gamma(1-\alpha)} \int_0^u \frac{f'(s)}{(u-s)^\alpha} ds$$

where, u might be one of the variables (x, y, z) .

Applications

The Fractional Geometry of a Parabola

Let \mathbf{r} be a parabola $t \rightarrow (t, t^2)$. Then we have

$$\mathbf{r}(t) = t\mathbf{e}_1 + t^2\mathbf{e}_2 \tag{59}$$

Hence:

$$\mathbf{r}_1(t) = \mathbf{e}_1 + \frac{{}^c D_t^a (t^2)}{{}^c D_t^a t} \mathbf{e}_2 = \mathbf{e}_1 + \frac{2t}{2-a} \mathbf{e}_2 \tag{60}$$

and

$$\mathbf{r}_2(t) = \frac{2}{2-a} \mathbf{e}_2$$

Then the centers of curvature of the parabola describe a curve:

$$\mathbf{c}(t) = c_1(t)\mathbf{e}_1 + c_2(t)\mathbf{e}_2 \tag{61}$$

Satisfying Eqs. 45 and 46 with

$$c_1 + \frac{2t}{2-a}c_2 = t + \frac{2t^3}{2-a} \tag{62}$$

$$\frac{2}{2-a}c_2 = \frac{2t^2}{2-a} + 1 + \frac{4t^2}{(2-a)^2} \tag{63}$$

Solving the system of Eqs. 62 and 63 we get:

$$c_1 = -\frac{4t^3}{(-2+a)^2} \tag{64}$$

$$c_2 = -\frac{4+a^2+8t^2-2a(2+t^2)}{(4-2a)} \tag{65}$$

Figure 3 shows the tangent space of the parabola at the point $t = 1.5$ for various values of the fractional dimension $\alpha = (1, 0.7, 0.3)$.

It is clear that the tangent spaces for $\alpha = 0.7$ and $\alpha = 0.3$ intersect the parabola at the point $t = 1.5$, although the conventional tangent space with fractional dimension $\alpha = 1.0$ touches the parabola at $t = 1.5$.

Furthermore the centers of curvature for various values of the fractional dimension α the point $t = 1,5$ are (for the conventional case):

$\alpha = 1.0$	$c_1 = -13.5$	and $c_2 = 7.25$
$\alpha = 0.7$	$c_1 = -7.98$	and $c_2 = 6.36$
$\alpha = 0.3$	$c_1 = -4.67$	and $c_2 = 5.75$

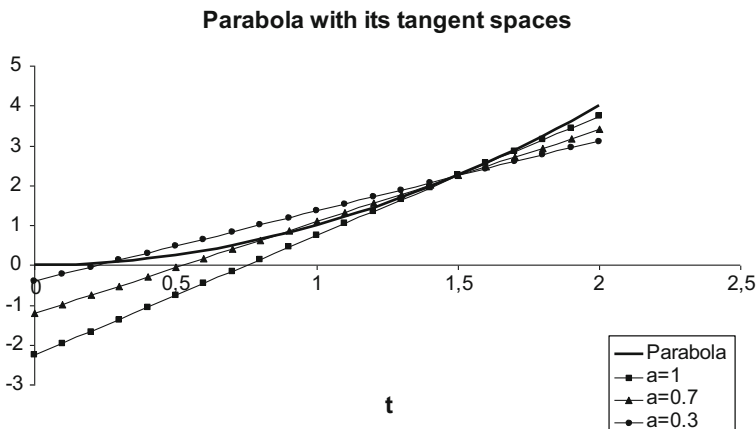


Fig. 3 The parabola with its tangent spaces at $t = 1.5$ for $\alpha = 1, \alpha = 0.7, \alpha = 0.3$

The Tangent and Curvature Center of the Weierstrass Function

Let us consider the function,

$$W(t) = \sum_{n=1}^{\infty} \lambda^{-\alpha n} \left\{ \sin\left(\frac{\lambda^n t}{2}\right) - \sin(2\lambda^n t) \right\} \tag{66}$$

the well-known Weierstrass function, continuous with discontinuous conventional derivatives at any point, (Liang and Su 2007). The parameter α has been proved to be related to the fractional dimension of the function $W(t)$. Restricting the function to $w(t)$ with

$$w(t) = \sum_{n=1}^6 \lambda^{-\alpha n} \left\{ \sin\left(\frac{\lambda^n t}{2}\right) - \sin(2\lambda^n t) \right\} \tag{67}$$

and for $\alpha = 0.5$ and $\lambda = 2$, the fractional tangent to the curve at the point $t = 1.0$ has been drawn, (Fig. 4), with the help of the Mathematica computerized pack.

Bending of Fractional Beams

Considering the pure fractional bending problem of a beam with microcracks, microvoids, various other defects, we get the fractional strain:

$$\varepsilon_{xx}^{\alpha} = -\frac{y}{\rho^{\alpha}} \tag{68}$$

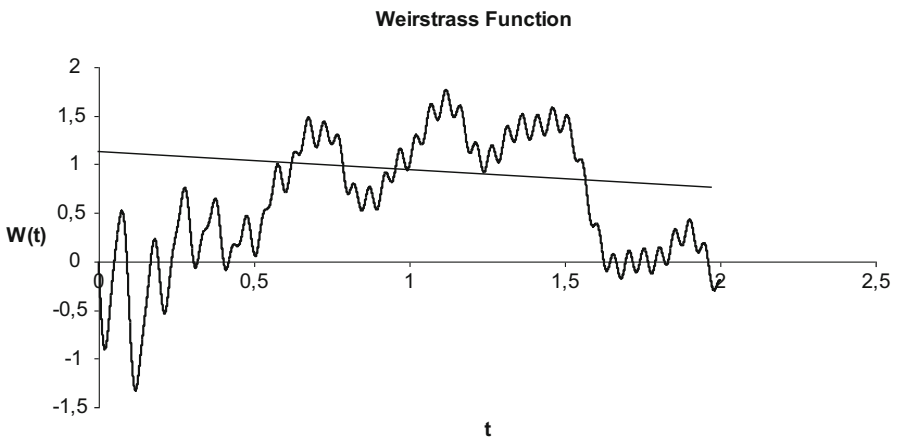


Fig. 4 The function $w(t)$ with its fractional ($\alpha = 0.5$) tangent at $t = 1.0$

where the fractional curvature is defined by,

$$\frac{1}{\rho^\alpha} = D_2 w(x) = \frac{1}{{}_a D_x^\alpha} {}_a D_x^\alpha \left(\frac{{}_a D_x^\alpha w(x)}{{}_a D_x^\alpha x} \right), \tag{69}$$

with $w(x)$ denoting the elastic line of the beam.

Likewise, the fractional bending moment is expressed by:

$$M = -2 \int_0^{h/2} \sigma_{xx}^\alpha y \, d^\alpha y = \frac{EI^\alpha}{\rho^\alpha} \tag{70}$$

with the fractional stress, (see Lazopoulos et al. 2015),

$$\sigma_{xx}^\alpha = -\frac{M}{I^\alpha} y \tag{71}$$

Hence the fractional bending of beams formula is revisited and expressed by:

$$M = EI^\alpha D_2 w(x) = \frac{EI^\alpha}{{}_0^c D_x^\alpha} D_x^\alpha \left(\frac{{}_0^c D_x^\alpha w(x)}{{}_0^c D_x^\alpha x} \right) \tag{72}$$

Therefore, the deflection curve $w(x)$ is defined by:

$$w(x) = \int_0^x \left(\int_0^s \frac{M(t)}{EI^\alpha} d^\alpha t \right) d^\alpha s + c_1 x + c_2 \tag{73}$$

In conventional integration, the deflection curve is defined by:

$$w(x) = \int_0^x \frac{s^{1-\alpha}}{\Gamma(2-\alpha)} \left(\int_0^s \frac{M(t)}{EI^\alpha} \frac{t^{1-\alpha}}{\Gamma(2-\alpha)} \frac{1}{\Gamma(\alpha)(s-t)^{1-\alpha}} dt \right) \frac{ds}{(x-s)^{1-\alpha}} + c_1 x + c_2 \tag{74}$$

The Fractional Tangent Plane of a Surface

Let us consider a manifold, with points $M(u,v)$, defined by the vectors

$$M(u, v) = \mathbf{x}(u, v) \tag{75}$$

with,

$$x_i = x_i(u, v), u_1 \leq u \leq u_2, v_1 \leq v \leq v_2, i = 1, 2, 3 \tag{76}$$

The infinitesimal distance between two points P and Q on the manifold M is defined by,

$$d^\alpha \mathbf{x} = \frac{{}^c D_u^\alpha \mathbf{x}}{{}^c D_u^\alpha u} d^\alpha u + \frac{{}^c D_v^\alpha \mathbf{x}}{{}^c D_v^\alpha v} d^\alpha v \tag{77}$$

In fact for the surface

$$z = u^2 v^2 \tag{78}$$

see, Fig. 2, the tangent space according to Eq. 77 is expressed by:

$$d^\alpha \mathbf{r} = d^\alpha x \mathbf{i} + d^\alpha y \mathbf{j} + \frac{2xy}{(2-\alpha)} (y d^\alpha x + x d^\alpha y) \mathbf{k} \tag{79}$$

Figures 5 and 6 shows the surface defined by Eq. 78 with its fractional tangent

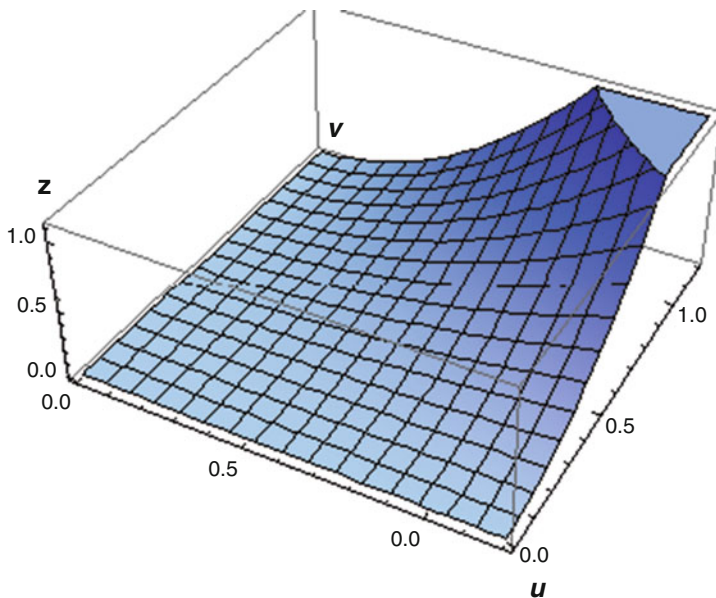


Fig. 5 The surface $z = u^2 v^2$ plane (space) at the point $(u,v) = (0.5, 0.5)$ for two fractional dimensions, $\alpha = 1$ (the conventional case) and $\alpha = 0.3$. It is clear that the fractional tangent plane is different from the conventional one ($\alpha = 1$).

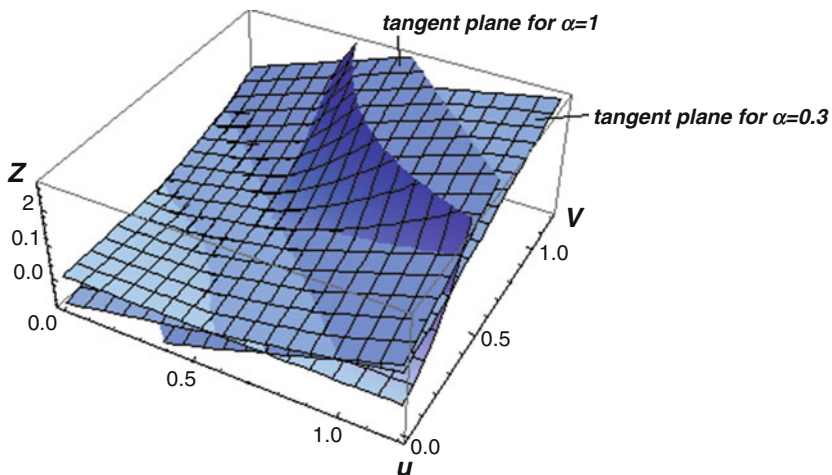


Fig. 6 The tangent planes for various values of the fractional dimension α

Fundamental Differential Forms on Fractional Differential Manifolds

The First Fractional Fundamental Form

Following formal procedure (Guggenheimer 1977), the quantity

$$\begin{aligned}
 I^\alpha &= d^\alpha \mathbf{x} \cdot d^\alpha \mathbf{x} = \left(\frac{{}^c D_u^\alpha \mathbf{x}}{{}^c D_u^\alpha u} d^\alpha u + \frac{{}^c D_v^\alpha \mathbf{x}}{{}^c D_v^\alpha v} d^\alpha v \right) \cdot \left(\frac{{}^c D_u^\alpha \mathbf{x}}{{}^c D_u^\alpha u} d^\alpha u + \frac{{}^c D_v^\alpha \mathbf{x}}{{}^c D_v^\alpha v} d^\alpha v \right) \\
 &= E d^\alpha u^2 + 2F d^\alpha u d^\alpha v + G d^\alpha v^2
 \end{aligned} \tag{80}$$

defined upon the tangent space of the manifold, as it has been clarified earlier, the I^α stands for the first fractional differential form, with the dot meaning the inner product.

$$\begin{aligned}
 E &= \frac{{}^c D_u^\alpha \mathbf{x}}{{}^c D_u^\alpha u} \cdot \frac{{}^c D_u^\alpha \mathbf{x}}{{}^c D_u^\alpha u} \\
 F &= \frac{{}^c D_u^\alpha \mathbf{x}}{{}^c D_u^\alpha u} \cdot \frac{{}^c D_v^\alpha \mathbf{x}}{{}^c D_v^\alpha v} \\
 G &= \frac{{}^c D_v^\alpha \mathbf{x}}{{}^c D_v^\alpha v} \cdot \frac{{}^c D_v^\alpha \mathbf{x}}{{}^c D_v^\alpha v}
 \end{aligned} \tag{81}$$

corresponding to

$$I^\alpha = E du^2 + 2F du dv + G dv^2$$

Furthermore the first fundamental form is positive definite, i.e., $0 \leq I^\alpha$ with $I^\alpha = 0$ if and only if $d^\alpha u$ and $d^\alpha v$ are equal to zero. Hence,

$$EG - F^2 > 0.$$

The Second Fractional Fundamental Form

Consider the manifold $M(u, v) = \mathbf{x}(u, v)$. Then, at each point of the manifold, there is a fractional unit normal \mathbf{N} to the fractional tangent plane

$$\mathbf{N} = \frac{\frac{{}_a D_u^\alpha \mathbf{x}}{{}_a D_u^\alpha u} \times \frac{{}_a D_v^\alpha \mathbf{x}}{{}_a D_v^\alpha v}}{\left| \frac{{}_a D_u^\alpha \mathbf{x}}{{}_a D_u^\alpha u} \times \frac{{}_a D_v^\alpha \mathbf{x}}{{}_a D_v^\alpha v} \right|}} \quad (82)$$

that is a function of u and v with the fractional differential

$$d^\alpha \mathbf{N} = \frac{{}_a D_u^\alpha \mathbf{N}}{{}_a D_u^\alpha u} d^\alpha u + \frac{{}_a D_v^\alpha \mathbf{N}}{{}_a D_v^\alpha v} d^\alpha v \quad (83)$$

Restricting only to Caputo's fractional derivatives with the property of zero fractional derivative of any constant, and taking into consideration that $\mathbf{N} \cdot \mathbf{N} = 1$, we get,

$$d^\alpha \mathbf{N} \cdot \mathbf{N} = 0 \quad (84)$$

where the vector $d^\alpha \mathbf{N}$ is parallel to the fractional tangent space.

The second fractional fundamental form is defined by Guggenheimer (1977),

$$\begin{aligned} \Pi^\alpha &= -d^\alpha \mathbf{x} \cdot d^\alpha \mathbf{N} = - \left(\frac{{}_a D_u^\alpha \mathbf{x}}{{}_a D_u^\alpha u} d^\alpha u + \frac{{}_a D_v^\alpha \mathbf{x}}{{}_a D_v^\alpha v} d^\alpha v \right) \cdot \left(\frac{{}_a D_u^\alpha \mathbf{N}}{{}_a D_u^\alpha u} d^\alpha u + \frac{{}_a D_v^\alpha \mathbf{N}}{{}_a D_v^\alpha v} d^\alpha v \right) \\ &= L d^\alpha u^2 + 2M d^\alpha u d^\alpha v + N d^\alpha v^2 \end{aligned} \quad (85)$$

with,

$$\begin{aligned} L &= - \frac{{}_a D_u^\alpha \mathbf{x}}{{}_a D_u^\alpha u} \cdot \frac{{}_a D_u^\alpha \mathbf{N}}{{}_a D_u^\alpha u} \\ M &= - \frac{1}{2} \left(\frac{{}_a D_u^\alpha \mathbf{x}}{{}_a D_u^\alpha u} \cdot \frac{{}_a D_v^\alpha \mathbf{N}}{{}_a D_v^\alpha v} + \frac{{}_a D_u^\alpha \mathbf{N}}{{}_a D_u^\alpha u} \cdot \frac{{}_a D_v^\alpha \mathbf{x}}{{}_a D_v^\alpha v} \right) \\ N &= - \frac{{}_a D_v^\alpha \mathbf{x}}{{}_a D_v^\alpha v} \cdot \frac{{}_a D_v^\alpha \mathbf{N}}{{}_a D_v^\alpha v} \end{aligned} \quad (86)$$

It is pointed again that the geometric procedures, that use quantities not defined upon the correct tangent spaces, are questionable. Even if analytically may yield the same results, geometrically are confusing.

The Fractional Normal Curvature

Let P be a point on a surface $\mathbf{x} = \mathbf{x}(u, v)$ and $\mathbf{x}(t) = \mathbf{x}(u(t), v(t))$ a regular curve C at P . The fractional curvature of curves has been discussed in ► Chap. 6, “Uniqueness of Elastoplastic Properties Measured by Instrumented Indentation”. The normal curvature k_n^α vector of C at P is the vector projection of the curvature vector \mathbf{k}^α onto the normal vector \mathbf{N} at P . The component of \mathbf{k}^α in the direction of the normal \mathbf{N} is called the normal fractional curvature of C at P and is denoted by k_n^α . Therefore,

$$k_n^\alpha = \mathbf{k}^\alpha \cdot \mathbf{N} \quad (87)$$

Since the unit tangent to C at P is the vector

$$\mathbf{t} = \frac{d^\alpha \mathbf{x}}{d^\alpha s} = \frac{d^\alpha \mathbf{x}}{d^\alpha t} / \left| \frac{d^\alpha \mathbf{x}}{d^\alpha t} \right| \quad (88)$$

where s denotes the fractional arc length of the curve and \mathbf{t} is the unit perpendicular to the normal \mathbf{N} along the curve, we get:

$$0 = \frac{d^\alpha (\mathbf{t} \cdot \mathbf{N})}{d^\alpha t} = \frac{d^\alpha \mathbf{t}}{d^\alpha t} \cdot \mathbf{N} + \mathbf{t} \cdot \frac{d^\alpha \mathbf{N}}{d^\alpha t} \quad (89)$$

Therefore, the normal curvature of a curve is equal to:

$$\begin{aligned} k_n^\alpha &= \mathbf{k} \cdot \mathbf{N} = \frac{d^\alpha \mathbf{t}}{d^\alpha t} \cdot \mathbf{N} / \left| \frac{d^\alpha \mathbf{x}}{d^\alpha t} \right| = -\mathbf{t} \cdot \frac{d^\alpha \mathbf{N}}{d^\alpha t} / \left| \frac{d^\alpha \mathbf{x}}{d^\alpha t} \right| \\ &= -\frac{d^\alpha \mathbf{x}}{d^\alpha t} \cdot \frac{d^\alpha \mathbf{N}}{d^\alpha t} / \left| \frac{d^\alpha \mathbf{x}}{d^\alpha t} \right|^2 \\ &= -\frac{\left(\frac{{}^c D_u^\alpha \mathbf{x}}{{}^c D_u^\alpha u} \frac{d^\alpha u}{d^\alpha t} + \frac{{}^c D_v^\alpha \mathbf{x}}{{}^c D_v^\alpha v} \frac{d^\alpha v}{d^\alpha t} \right) \cdot \left(\frac{{}^c D_u^\alpha \mathbf{N}}{{}^c D_u^\alpha u} \frac{d^\alpha u}{d^\alpha t} + \frac{{}^c D_v^\alpha \mathbf{N}}{{}^c D_v^\alpha v} \frac{d^\alpha v}{d^\alpha t} \right)}{\left(\frac{{}^c D_u^\alpha \mathbf{x}}{{}^c D_u^\alpha u} \frac{d^\alpha u}{d^\alpha t} + \frac{{}^c D_v^\alpha \mathbf{x}}{{}^c D_v^\alpha v} \frac{d^\alpha v}{d^\alpha t} \right) \cdot \left(\frac{{}^c D_u^\alpha \mathbf{x}}{{}^c D_u^\alpha u} \frac{d^\alpha u}{d^\alpha t} + \frac{{}^c D_v^\alpha \mathbf{x}}{{}^c D_v^\alpha v} \frac{d^\alpha v}{d^\alpha t} \right)} \\ &= \frac{L(d^\alpha u/d^\alpha t)^2 + 2M(d^\alpha u/d^\alpha t)(d^\alpha v/d^\alpha t) + N(d^\alpha v/d^\alpha t)^2}{E(d^\alpha u/d^\alpha t)^2 + 2F(d^\alpha u/d^\alpha t)(d^\alpha v/d^\alpha t) + G(d^\alpha v/d^\alpha t)^2} \quad (90) \end{aligned}$$

Recalling Eqs. 87 and 90, the normal curvature is defined by:

$$k_n^\alpha = \frac{\Pi^\alpha}{I^\alpha} \tag{91}$$

Fractional Vector Operators

In the present section the fractional tangent spaces along with their fractional normal vectors should be reminded, as they were defined in the preceding sections “The Geometry of Fractional Differential,” “The Fractional Arc Length,” and “The Fractional Tangent Space.”

For Cartesian coordinates, fractional generalizations of the divergence or gradient operators are defined by:

$$\nabla^{(a)} f(\mathbf{x}) = \text{grad}^{(a)} f(\mathbf{x}) = \nabla_i^{(a)} f(x) \mathbf{e}_i = \frac{{}_\omega^c D_i^a f(\mathbf{x})}{{}_\omega^c D_i^a x_i} + \mathbf{e}_i = {}_\omega^L D_i^a f(x) \mathbf{e}_i \tag{92}$$

where ${}_\omega^c D_i^a$ are Caputo’s fractional derivatives of order α and the subline meaning no contraction. Further, ${}_\omega^L D_i^a f(\mathbf{x})$ is Leibniz’s derivative, Eq. 18. Hence, the gradient of the vector \mathbf{x} is

$$\nabla^{(a)} \mathbf{x} = \mathbf{I} \tag{93}$$

with \mathbf{I} denoting the identity matrix. Consequently for a vector field

$$\mathbf{F}(x_1, x_2, x_3) = \mathbf{e}_1 F_1(x_1, x_2, x_3) + \mathbf{e}_2 F_2(x_1, x_2, x_3) + \mathbf{e}_3 F_3(x_1, x_2, x_3) \tag{94}$$

where $F_i(x_1, x_2, x_3)$ are absolutely integrable, the circulation is defined by:

$$C_L^{(a)}(\mathbf{F}) = \left({}_\omega I_L^{(a)}, \mathbf{F} \right) = \int_L (d\mathbf{L}, \mathbf{F}) = {}_\omega I^{(a)}_L (F_1 d^\alpha x_1) + {}_\omega I^{(a)}_L (F_2 d^\alpha x_2) + {}_\omega I^{(a)}_L (F_3 d^\alpha x_3) \tag{95}$$

Furthermore, the divergence of a vector $\mathbf{F}(\mathbf{x})$ is defined by:

$$\nabla^{(a)} \cdot \mathbf{F}(x) = \text{div}^{(a)} \mathbf{F}(x) = \frac{{}_\omega^c D_k^a F_k(\mathbf{x})}{{}_\omega^c D_k^a x_k} = {}_\omega^L D_k^a F_k(\mathbf{x}) \tag{96}$$

where the subline denotes no contraction.

Moreover, the fractional curl \mathbf{F} ($\text{curl}^{(a)}\mathbf{F}(\mathbf{x})$) of a vector \mathbf{F} is defined by:

$$\text{curl}^{(a)}\mathbf{F} = \mathbf{e}_l \varepsilon_{lmn} \frac{\omega^c D_m^a F_n}{\omega^c D_m^a x_m} = \mathbf{e}_l \varepsilon_{lmn\omega} {}^L D_m^a F_n \tag{97}$$

A fractional flux of the vector \mathbf{F} expressed in Cartesian coordinates across surface S is a fractional surface integral of the field with:

$$\Phi_s^\alpha(\mathbf{F}) = ({}_\omega I_s^\alpha, \mathbf{F}) = \iint_{\omega S}^{(\alpha)} (F_1 d^\alpha x_2 d^\alpha x_3 + F_2 d^\alpha x_3 d^\alpha x_1 + F_3 d^\alpha x_2 d^\alpha x_3) \tag{98}$$

A fractional volume integral of a triple fractional integral of a scalar field $f = f(x_1, x_2, x_3)$ is defined by:

$${}_\omega V_\Omega^{(a)} [f] = {}_\omega I_\Omega^{(a)} [x_1, x_2, x_3] f(x_1, x_2, x_3) = \iiint_{\omega \Omega}^{(\alpha)} f(x_1, x_2, x_3) d^\alpha x_1 d^\alpha x_2 d^\alpha x_3 \tag{99}$$

It should be pointed out that the triple fractional integral is not a volume integral, since the fractional derivative of a variable with respect to itself is different from one. So there is a clear distinction between the simple, double, or triple integrals and the line, surface, and volume integrals respectively.

Fractional Vector Field Theorems

Fractional Green's Formula

Green's theorem relates a line integral around a simple closed curve ∂B and a double integral over the plane region B with boundary ∂B . With positively oriented boundary ∂B , the conventional Greens theorem for a vector field $\mathbf{F} = F_1\mathbf{e}_1 + F_2\mathbf{e}_2$ is expressed by:

$$\int_{\partial B} (F_1 dx_1 + F_2 dx_2) = \int \int_B \left(\frac{\partial (F_1)}{\partial x_2} - \frac{\partial (F_2)}{\partial x_1} \right) dx_1 dx_2 \tag{100}$$

Recalling that:

$$d^a \mathbf{x} = (d^a x_1, d^a x_2) = ({}_\omega^c D_{x_1}^a [x_1] d^\alpha x_1, {}_\omega^c D_{x_2}^a [x_2] d^\alpha x_2) \tag{100a}$$

and substituting into conventional Green's theorem Eq. 100 we get:

$$\int_{\omega \partial W}^{(\alpha)} (F_1 d^\alpha x_1 + F_2 d^\alpha x_2) = \int_{\omega}^{(\alpha)} \int_W \left(\frac{c}{\omega} D_{x_2}^a (F_1) - \frac{c}{\omega} D_{x_1}^a (F_2) \right) d^\alpha x_1 d^\alpha x_2 \quad (101)$$

Fractional Stoke's Formula

Restricting in the consideration of a simple surface W , if we denote its boundary by ∂W and if \mathbf{F} is a vector field defined on W , then the conventional Stokes' Theorem asserts that:

$$\oint_W \mathbf{F} \cdot d\mathbf{L} = \iint_W \text{curl } \mathbf{F} \cdot d\mathbf{S} \quad (102)$$

In Cartesian coordinates it yields:

$$\begin{aligned} \int_{\partial W} (F_1 dx_1 + F_2 dx_2 + F_3 dx_3) &= \iint_W \left(\frac{\partial(F_3)}{\partial x_2} - \frac{\partial(F_2)}{\partial x_3} \right) dx_2 dx_3 \\ &+ \left(\frac{\partial(F_1)}{\partial x_3} - \frac{\partial(F_3)}{\partial x_1} \right) dx_3 dx_1 + \left(\frac{\partial(F_2)}{\partial x_1} - \frac{\partial(F_1)}{\partial x_2} \right) dx_1 dx_2 \end{aligned} \quad (103)$$

where $\mathbf{F}(x_1, x_2, x_3) = \mathbf{e}_1 F_1(x_1, x_2, x_3) + \mathbf{e}_2 F_2(x_1, x_2, x_3) + \mathbf{e}_3 F_3(x_1, x_2, x_3)$.

In this case, the fractional curl operation is defined by:

$$\begin{aligned} \text{curl}_W^\alpha (\mathbf{F}) &= e_l \varepsilon_{lmn} \frac{\alpha}{\omega} D_{x_m}^a (F_n) / \frac{c}{\omega} D_{x_k}^a (x_i) \delta_{mk} = \mathbf{e}_1 \left(\frac{\alpha}{\omega} D_{x_2}^a F_3 - \frac{\alpha}{\omega} D_{x_3}^a F_2 \right) \\ &+ \mathbf{e}_2 \left(\frac{\alpha}{\omega} D_{x_3}^a F_1 - \frac{\alpha}{\omega} D_{x_1}^a F_3 \right) + \mathbf{e}_3 \left(\frac{\alpha}{\omega} D_{x_1}^a F_2 - \frac{\alpha}{\omega} D_{x_2}^a F_1 \right) \end{aligned} \quad (104)$$

Therefore transforming the conventional Stokes' theorem into the fractional form we get:

$$\begin{aligned} \int_{\omega}^{(\alpha)} \int_W (F_1 d^\alpha x_1 + F_2 d^\alpha x_2 + F_3 d^\alpha x_3) &= \int_{\omega}^{(\alpha)} \iint_W \left\{ \left(\frac{\alpha}{\omega} D_{x_2}^a F_3 - \frac{\alpha}{\omega} D_{x_3}^a F_2 \right) d^\alpha x_2 d^\alpha x_3 \right. \\ &+ \left(\frac{\alpha}{\omega} D_{x_3}^a F_1 - \frac{\alpha}{\omega} D_{x_1}^a F_3 \right) d^\alpha x_3 d^\alpha x_1 \\ &+ \left. \left(\frac{\alpha}{\omega} D_{x_1}^a F_2 - \frac{\alpha}{\omega} D_{x_2}^a F_1 \right) d^\alpha x_1 d^\alpha x_2 \right\} \end{aligned} \quad (105)$$

Fractional Gauss' Formula

For the conventional fields theory, let $\mathbf{F} = \mathbf{e}_1 F_1 + \mathbf{e}_2 F_2 + \mathbf{e}_3 F_3$, be a continuously differentiable real-valued function in a domain W with boundary ∂W . Then the conventional divergence Gauss' theorem is expressed by:

$$\iint_{\partial W} \mathbf{F} \cdot \mathbf{dS} = \iiint_W \text{div} \mathbf{F} dV \quad (106)$$

Since

$$\mathbf{d}^{(\alpha)} \mathbf{S} = \mathbf{e}_1 d^\alpha x_2 d^\alpha x_3 + \mathbf{e}_2 d^\alpha x_3 d^\alpha x_1 + \mathbf{e}_3 d^\alpha x_1 d^\alpha x_2 \quad (107)$$

where $d^\alpha x_i$, $i = 1, 2, 3$ is expressed by Eq. 15,

$$d^{(\alpha)} V = d^\alpha x_1 d^\alpha x_2 d^\alpha x_3 \quad (108)$$

Furthermore, see Eq. 96,

$$\text{div}^{(\alpha)} \mathbf{F}(x) = \frac{{}^c D_k^\alpha F_k(\mathbf{x})}{{}^c D_k^\alpha x_i} \delta_{km} \quad (109)$$

The Fractional Gauss divergence theorem becomes:

$$\iint_{\partial W} \mathbf{F} \cdot d^{(\alpha)} \mathbf{S} = \iiint_W \text{div}^{(\alpha)} \mathbf{F} d^{(\alpha)} V \quad (110)$$

Remember that the differential $d^\alpha S = \mathbf{n}^\alpha d^\alpha S$, where \mathbf{n}^α is the unit normal of the fractional tangent space as it has been defined in section “[The Geometry of Fractional Differential](#).”

Fractional Deformation Geometry

Outlining the Fractional Deformation Geometry presented in Lazopoulos and Lazopoulos (2016a), we assume the description in the Euclidean space, considering the reference configuration B with the boundary ∂B of a body displaced to its current configuration b with the boundary ∂b (see Truesdell 1977). The points in the reference placement B , defining the material points, are described by \mathbf{X} , whereas the set of the displaced points \mathbf{y} describe the current configuration b with the boundary ∂b . The coordinate system in B is denoted by $\{X_A\}$, while the corresponding to the current configuration b is reflected to $\{y_i\}$. In the present description, both systems have the same axial directions with base vectors $\{\mathbf{e}_A\}, \{\mathbf{e}_i\}$ whether the

reference concerns the current or the initial (unstressed) configuration. The motion of a reference point \mathbf{X} is described by the function

$$\mathbf{y} = \Psi(\mathbf{X}, t) \tag{111}$$

the conventional gradient of the deformation is defined by:

$$\mathbf{F}(\Psi, t) = \frac{\partial \Psi(\mathbf{X}, t)}{\partial \mathbf{X}} \text{ or } F_{iA} = \frac{\partial \Psi_i}{\partial X_A} \tag{112}$$

with

$$\mathbf{F}^{(\alpha)} = F_{ij}^{(\alpha)} = \nabla_{\mathbf{X}}^{(\alpha)} \mathbf{y}_i = {}^c D_{X_j}^\alpha \mathbf{y}_i = \frac{{}^c D_{X_j}^\alpha (\mathbf{X} + \mathbf{u})}{{}^c D_{X_j}^\alpha \mathbf{X}} = \nabla_{\mathbf{X}}^{(\alpha)} \mathbf{X} + \nabla_{\mathbf{X}}^{(\alpha)} \mathbf{u} \tag{113}$$

Furthermore, the right Cauchy-Green fractional deformation tensor,

$$\mathbf{C}^{(\alpha)} = \mathbf{F}^{(\alpha)T} \mathbf{F}^{(\alpha)} \tag{114}$$

and

$$\mathbf{B}^{(\alpha)} = \mathbf{F}^{(\alpha)} \mathbf{F}^{(\alpha)T} \tag{115}$$

is the left Cauchy-Green fractional deformation tensor.

Likewise, the fractional nonlinear fractional Green-Lagrange strain tensor $\mathbf{E}^{(\alpha)}$ may be defined by:

$$E_N^{(\alpha)} = \mathbf{N} \cdot \overset{\alpha}{\mathbf{E}} \mathbf{N} \tag{116}$$

where \mathbf{N} is the unit vector of the considered fiber in the reference placement, with

$$\mathbf{E}^{(\alpha)} = \frac{1}{2} (\mathbf{C}^{(\alpha)} - \mathbf{I}) \tag{117}$$

Recalling that the current placement $\mathbf{y} = \mathbf{X} + \mathbf{u}$, where \mathbf{u} denotes the displacement vector, the fractional Green-Lagrange deformation tensor becomes:

$$\mathbf{E}^{(\alpha)} = \frac{1}{2} \left[\left(\nabla_{\mathbf{X}}^{(\alpha)} \mathbf{u} \right)^T + \nabla_{\mathbf{X}}^{(\alpha)} \mathbf{u} + \left(\nabla_{\mathbf{X}}^{(\alpha)} \mathbf{u} \right)^T \nabla_{\mathbf{X}}^{(\alpha)} \mathbf{u} \right] \tag{118}$$

Proceeding to define the fractional strain tensor referred to the current placement, i.e., Euler-Almansi strain tensor:

$$\mathbf{A}^{(\alpha)} = \frac{1}{2} \left(\mathbf{I} - \left(\mathbf{B}^{(\alpha)} \right)^{-1} \right) \tag{119}$$

with the strain

$$\varepsilon_{\mathbf{n}} = \mathbf{n}^T \cdot \mathbf{A}^{(\alpha)} \mathbf{n} \quad (120)$$

where \mathbf{n} is the unit vector along the deformed fiber, corresponding to the \mathbf{N} unit vector along the reference placement fiber. It is evident that in the conventional case with $\alpha = 1$ the fractional Euler Almansi strain tensor $\mathbf{A}^{(\alpha)}$ reduces to the conventional strain tensor \mathbf{A} . It should be pointed out that the fractional stretches $\lambda^{(\alpha)}$, adopting the right Cauchy-Green strain tensor $\mathbf{C}^{(\alpha)}$, are defined by:

$$\lambda^{(\alpha)} = \frac{d^{(\alpha)}s}{d^{(\alpha)}S} \quad (121)$$

as the ratio of the measures of the final infinitesimal length over the corresponding length of the fractional differential $d^{(\alpha)}\mathbf{X}$ vector with $d^{(\alpha)}S = (d^{(\alpha)}\mathbf{X} \cdot d^{(\alpha)}\mathbf{X})^{\frac{1}{2}}$. In fact the stretches $\lambda^{(\alpha)}$ are defined by:

$$\left(\lambda^{(\alpha)}\right)^2 = \mathbf{N}^T \cdot \mathbf{C}^{(\alpha)} \mathbf{N} \quad (122)$$

where $\mathbf{N} = N_A e_A$ is the unit vector directed along the material reference fiber.

Furthermore for

$$\frac{1}{\left(\lambda^{(\alpha)}\right)^2} = \left(\frac{d^{(\alpha)}S}{d^{(\alpha)}s}\right)^2 = \mathbf{n}^T \cdot \left(\mathbf{B}^{(\alpha)}\right)^{-1} \mathbf{n} \quad (123)$$

where \mathbf{n} the unit vector along the deformed fiber.

Polar Decomposition of the Deformation Gradient

It is well known that every nonsingular matrix may be decomposed into a product of an orthogonal and a symmetric positive tensor. Applying the property to the deformation gradient we get:

$$\mathbf{F}^{(\alpha)} = \mathbf{R}^{(\alpha)} \mathbf{U}^{(\alpha)} = \mathbf{V}^{(\alpha)} \mathbf{R}^{(\alpha)} \quad (124)$$

where \mathbf{R} is orthogonal $\mathbf{R}=\mathbf{R}^{-T}$ and \mathbf{U} and \mathbf{V} are symmetric positive ($\mathbf{U}=\mathbf{U}^T$ and $\mathbf{V}=\mathbf{V}^T$). Therefore:

$$\mathbf{C}^{(\alpha)} = \left(\mathbf{U}^{(\alpha)}\right)^2 \text{ and } \mathbf{B}^{(\alpha)} = \left(\mathbf{V}^{(\alpha)}\right)^2 \quad (125)$$

Moreover, the eigenvalues of $\mathbf{C}^{(\alpha)}$ and $\mathbf{V}^{(\alpha)}$ are the same, but the eigenvectors $\mathbf{u}^{(\alpha)}$ of $\mathbf{U}^{(\alpha)}$ and $\mathbf{v}^{(\alpha)}$ of $\mathbf{V}^{(\alpha)}$ are related by $\mathbf{v}^{(\alpha)} = \mathbf{R}^{(\alpha)} \mathbf{u}^{(\alpha)}$. In fact $\mathbf{v}^{(\alpha)}$ is directed

along a principal direction (eigenvector) of the strain tensor $\mathbf{V}^{(\alpha)}$ with $\mathbf{u}^{(\alpha)}$ been the eigenvector of $\mathbf{U}^{(\alpha)}$. In other words, principal directions refer to the vectors $\mathbf{u}^{(\alpha)} = u_A^{(\alpha)} e_A$ and $\mathbf{v}^{(\alpha)} = v_A^{(\alpha)} e_A$.

Deformation of Volume and Surface

Consider three noncoplanar line elements $d^\alpha \mathbf{X}^{(1)}$, $d^\alpha \mathbf{X}^{(2)}$, $d^\alpha \mathbf{X}^{(3)}$ at the point \mathbf{X} in B so that:

$$\mathbf{d}^\alpha \mathbf{y}^{(i)} = \mathbf{F}^{(\alpha)} d^\alpha \mathbf{X}^{(i)} \quad (126)$$

with $d^\alpha \mathbf{y}^i$ the corresponding fractional differential vectors in the current placement.

Further, the volume $d^\alpha V$ is derived by

$$d^\alpha V = d^\alpha \mathbf{X}^{(1)} \cdot (d^\alpha \mathbf{X}^{(2)} \wedge d^\alpha \mathbf{X}^{(3)}) \quad (127)$$

Alternatively

$$d^\alpha V = \det(d^\alpha \mathbf{X}^{(1)}, d^\alpha \mathbf{X}^{(2)}, d^\alpha \mathbf{X}^{(3)}) \quad (128)$$

in which $d^\alpha \mathbf{X}^{(i)}$ denotes a column vector ($i = 1, 2, 3$). The corresponding volume $d^\alpha v$ in the deformed configuration is

$$d^\alpha v = \det(d^\alpha \mathbf{y}^{(1)}, d^\alpha \mathbf{y}^{(2)}, d^\alpha \mathbf{y}^{(3)}) \quad (129)$$

and

$$d^\alpha v = \det(\mathbf{F}^{(\alpha)}) d^\alpha V \equiv J^{(\alpha)} d^\alpha V \quad (130)$$

where

$$J^{(\alpha)} = \det \mathbf{F}^{(\alpha)} \quad \text{and} \quad d^\alpha V = d^\alpha \mathbf{X}^{(1)} \cdot d^\alpha \mathbf{X}^{(2)} \wedge d^\alpha \mathbf{X}^{(3)} \quad (131)$$

Consider, further, an infinitesimal vector element of material surface dS in the neighborhood of the point \mathbf{X} in B with $d^\alpha \mathbf{S} = \mathbf{N}^\alpha d^\alpha S$ the fractional tangent surface vector corresponding to the normal fractional normal vector \mathbf{N}^α . Likewise, $d^\alpha \mathbf{X}$ is an arbitrary fiber cutting the edge $d^\alpha S$ such that $d^\alpha \mathbf{X} \cdot d^\alpha \mathbf{S} > 0$. The volume of the cylinder with base $d^\alpha \mathbf{S}$ and generators $d^\alpha \mathbf{X}$ has volume $d^\alpha V = d^\alpha \mathbf{X} \cdot d^\alpha \mathbf{S}$. If $d^\alpha \mathbf{x}$ and $d^\alpha \mathbf{s}$ are the deformed configurations of $d^\alpha \mathbf{X}$ and $d^\alpha \mathbf{S}$ respectively, with $d^\alpha \mathbf{s} = \mathbf{n}^\alpha d^\alpha s$, where \mathbf{n}^α is the normal vector to the deformed surface, the volume $d^\alpha V$ in the reference placement corresponds to the volume $d^\alpha v = d^\alpha \mathbf{x} \cdot d^\alpha \mathbf{s}$ in the current configuration so that:

$$d^\alpha v = d^\alpha \mathbf{y} \cdot d^\alpha \mathbf{s} = J^{(\alpha)} d^\alpha \mathbf{X} \cdot d^\alpha \mathbf{S} \quad (132)$$

Since $d^\alpha \mathbf{y} = \mathbf{F}^{(\alpha)} d^\alpha \mathbf{X}$, we obtain

$$\mathbf{F}^{(\alpha)T} d^\alpha \mathbf{X} \cdot d^\alpha \mathbf{s} = J^{(\alpha)} d^\alpha \mathbf{X} \cdot d^\alpha \mathbf{S} \quad (133)$$

removing the arbitrary $d^\alpha \mathbf{X}$. Consequently,

$$d^\alpha \mathbf{s} = J^{(\alpha)} \left(\mathbf{F}^{(\alpha)} \right)^{-T} d^\alpha \mathbf{S} \quad (134)$$

and

$$\mathbf{n}^\alpha d^\alpha s = J^{(\alpha)} \left(\mathbf{F}^{(\alpha)} \right)^{-T} \mathbf{N}^\alpha d^\alpha S \quad (135)$$

The relation between the area elements corresponding to reference and current configurations is the well-known Nanson's formula for the fractional deformations.

Examples of Deformations

Homogeneous Deformations

The most general homogeneous deformation of the body B from its reference configuration is expressed by:

$$\mathbf{x} = \mathbf{A}\mathbf{X} \quad (136)$$

Choosing as fractional derivatives Caputo ones that are given by:

$${}_a^C D_t^\alpha (t-a)^v = \frac{\Gamma(v+1)}{\Gamma(-\alpha+v+1)} (t-a)^{v-\alpha}$$

and

$${}_a^C D_X^\alpha X = \frac{\Gamma(2)}{\Gamma(2-\alpha)} X^{1-\alpha} \quad (137)$$

and specializing the homogeneous deformations with the example of simple shear, we discuss the deformation,

$$\begin{aligned} x_1 &= X_1 + \gamma X_2 \\ x_2 &= X_2 \\ x_3 &= X_3 \end{aligned} \quad (138)$$

Therefore the Cauchy-Green deformation tensors $\mathbf{C}^{(\alpha)}$ and $\mathbf{B}^{(\alpha)}$ become:

$$\mathbf{C}^{(\alpha)} = \left(\mathbf{F}^{(\alpha)} \right)^T \mathbf{F}^{(\alpha)} = \begin{vmatrix} 1 & \gamma & 0 \\ \gamma & \gamma^2 + 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} \tag{139}$$

and

$$\mathbf{B} = \mathbf{F}^{(\alpha)} \left(\mathbf{F}^{(\alpha)} \right)^T = \begin{vmatrix} 1 + \gamma^2 & \gamma & 0 \\ \gamma & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} \tag{140}$$

The Green (Langrange) Fractional strain tensor is:

$$\mathbf{E}^{(\alpha)} = \frac{1}{2} \left(\mathbf{C}^{(\alpha)} - \mathbf{I} \right) = \begin{bmatrix} 0 & \gamma/2 & 0 \\ \gamma/2 & \gamma^2/2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{141}$$

And the Euler-Almansi strain tensor is given by:

$$\mathbf{A}^{(\alpha)} = \frac{1}{2} \left(\mathbf{I} - \left(\mathbf{B}^{(\alpha)} \right)^{-1} \right) = \begin{bmatrix} 0 & \gamma/2 & 0 \\ \gamma/2 & -\gamma^2/2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{142}$$

It would seem strange that the results are exactly the same as the ones of the conventional elasticity. However, there is mathematical explanation for the homogeneous deformations. Just taking into consideration the definition of the fractional derivative and differential, the differential for a linear function of the form:

$$f(x) = Ax \tag{143}$$

The fractional differential is given by:

$$d^\alpha f(x) = A d^\alpha x \tag{144}$$

the fractional differential of the function has almost the same form as the conventional one. Nevertheless, for the nonlinear function:

$$f(x) = x^4 \tag{145}$$

the fractional differential

$$d^\alpha f(x) = {}^C D_x^\alpha x^4 \cdot \left({}^C D_x^\alpha x \right)^{-1} d^\alpha x \tag{146}$$

Is equal to

$$d^\alpha f(x) = \frac{\Gamma(5)\Gamma(2-\alpha)}{\Gamma(5-\alpha)} x^3 d^\alpha x \quad (147)$$

with coefficient depending upon the α fractional dimension. That makes the difference in nonhomogeneous deformations discussed in the next section.

The Nonhomogeneous Deformations

The nonhomogeneous deformation is defined by the equations:

$$\begin{aligned} x_1 &= X_1 + \gamma X_2^4 \\ x_2 &= X_2 \\ x_3 &= X_3 \end{aligned} \quad (148)$$

Taking into consideration Eqs. 124 and 125, the fractional Cauchy-Green deformation tensors are expressed by:

$$\mathbf{C}^{(\alpha)} = \left(\mathbf{F}^{(\alpha)} \right)^T \mathbf{F}^{(\alpha)} = \begin{vmatrix} 1 & \frac{24\gamma\Gamma(2-\alpha)X_2^3}{\Gamma(5-\alpha)} & 0 \\ \frac{24\gamma\Gamma(2-\alpha)X_2^3}{\Gamma(5-\alpha)} & 1 + \frac{576\gamma^2\Gamma(2-\alpha)^2X_2^6}{\Gamma(5-\alpha)^2} & 0 \\ 0 & 0 & 1 \end{vmatrix} \quad (149)$$

and

$$\mathbf{B}^{(\alpha)} = \mathbf{F}^{(\alpha)} \left(\mathbf{F}^{(\alpha)} \right)^T = \begin{vmatrix} 1 + \frac{576\gamma^2\Gamma(2-\alpha)^2X_2^6}{\Gamma(5-\alpha)^2} & \frac{24\gamma\Gamma(2-\alpha)X_2^3}{\Gamma(5-\alpha)} & 0 \\ \frac{24\gamma\Gamma(2-\alpha)X_2^3}{\Gamma(5-\alpha)} & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} \quad (150)$$

Hence Green-Lagrange strain tensor is expressed by:

$$\mathbf{E}^{(\alpha)} = \begin{vmatrix} 0 & \frac{12\gamma\Gamma(2-\alpha)X_2^3}{\Gamma(5-\alpha)} & 0 \\ \frac{12\gamma\Gamma(2-\alpha)X_2^3}{\Gamma(5-\alpha)} & \frac{288\gamma^2\Gamma(2-\alpha)^2X_2^6}{\Gamma(5-\alpha)^2} & 0 \\ 0 & 0 & 0 \end{vmatrix} \quad (151)$$

and Euler-Almansi strain tensor is defined by:

$$\mathbf{A}^{(\alpha)} = \begin{vmatrix} 0 & \frac{12\gamma\Gamma(2-\alpha)X_2^3}{\Gamma(5-\alpha)} & 0 \\ \frac{12\gamma\Gamma(2-\alpha)X_2^3}{\Gamma(5-\alpha)} & -\frac{288\gamma^2\Gamma(2-\alpha)^2X_2^6}{\Gamma(5-\alpha)^2} & 0 \\ 0 & 0 & 0 \end{vmatrix} \quad (152)$$

It is evident that the strain tensors strongly depend upon the fractional dimension α for the present nonhomogeneous deformation. However, the present deformation also depends upon the source ω of the fractional analysis.

The Infinitesimal Deformations

Since there has been pointed out in the introduction, fractional strain tensors have been considered in the literature, mainly in infinitesimal deformations, simply by substituting the common derivatives to fractional ones. It would be wise to study whether that idea is valid or not. Unfortunately it is proven a mistake. Fractional strain with simple substitution of derivatives does not have any physical meaning.

Considering the fractional Euler-Lagrange strain tensor, Eq. 118, where \mathbf{u} is the small displacement vector with $|\mathbf{u}| \ll 1$ and $|\nabla_{\mathbf{x}}^{(\alpha)} \mathbf{u}| \ll 1$, we restrict into the linear deformation analysis, and the infinitesimal (linear) fractional Euler-Lagrange strain tensor $\mathbf{E}_{lin}^{\alpha}$ becomes:

$$\mathbf{E}_{lin}^{(\alpha)} = \frac{1}{2} \left[(\nabla_{\mathbf{x}}) \mathbf{u}^T + \nabla_{\mathbf{x}} \mathbf{u} \right] \quad (153)$$

It is recalled from Eq. 92 that:

$$\nabla^{(\alpha)} f(\mathbf{x}) = grad^{(\alpha)} f(\mathbf{x}) = \nabla_i^{(\alpha)} f(x) \mathbf{e}_i = \frac{{}^c D_i^a f(\mathbf{x})}{{}^c D_i^a x_i} \mathbf{e}_i, \quad (154)$$

Hence the linear fractional strain is not simply the half of the sum of the fractional derivatives of the displacement vector and its transport, but the half of the fractional gradient (as it has been defined by Eq. 92) and its transport.

Fractional Stresses

Pointing out that the fractional tangent space of a surface has different orientation of the conventional one, the fractional normal vector \mathbf{n}^{α} does not coincide with the conventional normal vector \mathbf{n} . Hence we should expect the stresses and consequently the stress tensor to differ from the conventional ones not only in the values.

If $d^{\alpha} \mathbf{P}$ is a contact force acting on the deformed area $d^{\alpha} \mathbf{a} = \mathbf{n}^{\alpha} d^{\alpha} a$, lying on the fractional tangent plane where \mathbf{n}^{α} is the unit outer normal to the element of area $d^{\alpha} \mathbf{a}$, then the α -fractional stress vector is defined by:

$$\mathbf{t}^{\alpha} = \lim_{d^{\alpha} a \geq 0} \frac{d^{\alpha} \mathbf{P}}{d^{\alpha} a} \quad (155)$$

However, the α -fractional stress vector does not have any connection with the conventional one

$$\mathbf{t} = \lim_{da \geq 0} \frac{d\mathbf{P}}{da} \quad (156)$$

since the conventional tangent plane has different orientation from the α -fractional tangent plane and the corresponding normal vectors too.

Following similar procedures as the conventional ones we may establish Cauchy's fundamental theorem (see Truesdell 1977).

If $\mathbf{t}^\alpha(\cdot, \mathbf{n}^\alpha)$ is a continuous function of the transplacement vector \mathbf{x} , there is an α -fractional Cauchy stress tensor field

$$\mathbf{T}^\alpha = \left[\sigma_{ij}^\alpha \right] \quad (157)$$

The Balance Principles

Almost all balance principles are based upon Reynold's transport theorem. Hence the modification of that theorem, just to conform to fractional analysis is presented. The conventional Reynold's transport theorem is expressed by:

$$\frac{d}{dt} \int_W A dV = \int_W \frac{dA}{dt} dV + \int_{\partial W} A \mathbf{v}_n dS \quad (158)$$

For a vector field A applied upon region W with boundary ∂W and \mathbf{v}_n is the normal velocity of the boundary ∂W .

Material Derivatives of Volume, Surface, and Line Integrals

For any scalar, vector or tensor property that may be represented by:

$$P_{ij}(t) = \int_V P_{ij}^*(x, t) dV \quad (159)$$

where V is the volume of the current placement in the conventional calculus. The material time derivative of $P_{ij}^*(x, t)$ is expressed by:

$$\frac{dP_{ij}^*}{dt} = \frac{\partial P_{ij}^*}{\partial t} + v \frac{\partial P_{ij}^*}{\partial x_j} = \frac{\partial P_{ij}^*}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{x}} P_{ij}^* \quad (160)$$

recalling that we consider constant material points during time derivation. Similarly in fractional calculus, the material time derivative is given, for any tensor field P_{ij} by:

$$\frac{dP_{ij}^*}{dt} = \frac{\partial P_{ij}^*}{\partial t} + \frac{\partial^\alpha P_{ij}^*}{\partial x_k^a} \frac{d^a x_l}{dx_k^a} \frac{dg(x_m)}{dt} \delta_{km} \tag{161}$$

Since

$$v_k = \frac{\partial^a x_e}{\partial x_k^a} \frac{dg(x_m)}{dt} \delta_{km} \tag{162}$$

The material derivative, into the context of fractional calculus, is expressed by:

$$\frac{d(\quad)}{dt} = \frac{\partial(\quad)}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{x}}^{(\alpha)} \tag{163}$$

Hence, the acceleration is defined by:

$$\mathbf{a} = \frac{d\mathbf{v}}{dt} = \frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{x}}^{(\alpha)} \mathbf{v} \tag{164}$$

Furthermore the material time derivative of $P_{ij}(t)$ is expressed in conventional analysis by:

$$\frac{d}{dt} [P_{ij}(t)] = \frac{d}{dt} \int_V P_{ij}^*(x, t) dV \tag{165}$$

where b is the current placement of the region. It is well known that:

$(d\dot{V}) = JdV$, and the Eq. 165 yields:

$$\frac{d}{dt} \int_V P_{ij}^*(x, t) dV = \int_V \left[\frac{\partial P_{ij}^*(x, t)}{\partial t} + P_{ij}^*(x, t) \frac{\partial v_p}{\partial x_p} \right] dV \tag{166}$$

Recalling the material derivative operator Eqs. 160 and 166 yields,

$$\frac{d}{dt} \int_V P_{ij}^*(x, t) dV = \int_b \left[\frac{\partial P_{ij}^*(x, t)}{\partial t} + \frac{\partial (u_p P_{ij}^*(x, t))}{\partial x_p} \right] dV \tag{167}$$

yielding Reynold's Transport theorem:

$$\frac{d}{dt} \int_v P_{ij}^*(x, t) dV = \int_V \frac{\partial P_{ij}^*(x, t)}{\partial t} dV + \int_S v_p [P_{ij}^*(x, t)] dS_p \quad (168)$$

Expressing Reynold's Transfer Theorem in fractional form we get:

$$\frac{d}{dt} \int_{\omega V} P_{ij}^*(x, t) d^\alpha V = \int_{\omega V} \frac{\partial P_{ij}^*(x, t)}{\partial t} d^\alpha V + \int_{\omega S} u_p [P_{ij}^*(x, t)] d^\alpha S_p \quad (169)$$

where $d^\alpha V$ and $d^\alpha S$ the infinitesimal fractional volume and surface, respectively.

The volume integral of the material time derivative of $P_{ij}(t)$ may also be expressed by:

$$\frac{d}{dt} \int_{\omega V}^{(\alpha)} P_{ij}^*(\mathbf{x}, t) d^\alpha V = \int_{\omega V}^{(\alpha)} \left(\frac{\partial P_{ij}^*(x, t)}{\partial t} + div^\alpha [\mathbf{v} P_{ij}^*] \right) d^\alpha V \quad (170)$$

The Balance of Mass

The conventional balance of mass, expressing the mass preservation, is expressed by:

$$\frac{d}{dt} \int_W \rho dV = 0 \quad (171)$$

In the fractional form it is given by:

$$\frac{d}{dt} \int_{\omega W}^{(\alpha)} \rho d^\alpha V = 0 \quad (172)$$

where $\frac{d}{dt}$ is the total time derivative.

Recalling the fractional Reynold's Transport Theorem, we get:

$$\frac{d}{dt} \int_{\omega V}^{(\alpha)} \rho d^\alpha V = \int_{\omega V}^{(\alpha)} \left(\frac{\partial \rho(x, t)}{\partial t} + div^\alpha [\mathbf{v} \rho] \right) d^\alpha V \quad (173)$$

Since Eq. 173 is valid for any volume V , the continuity equation is:

$$\frac{\partial \rho}{\partial t} + \operatorname{div}^a [\mathbf{v}\rho] = 0 \quad (174)$$

where div^a has already been defined by Eq. 96. That is the continuity equation expressed in fractional form.

Balance of Linear Momentum Principle

It is reminded that the conventional balance of linear momentum is expressed in continuum mechanics by:

$$\frac{d}{dt} \int_{\Omega} \rho \mathbf{v} dV = \int_{\partial\Omega} \mathbf{t}^{(n)} dS + \int_{\Omega} \rho \mathbf{b} dV \quad (175)$$

where \mathbf{v} is the velocity, $\mathbf{t}^{(n)}$ is the traction on the boundary, and \mathbf{b} is the body force per unit mass. Likewise that principle in fractional form is expressed by:

$$\frac{d}{dt} \int_{\omega \Omega}^{(\alpha)} \rho \mathbf{v} d^{(\alpha)} V = \int_{\omega \Omega}^{(\alpha)} [\rho \mathbf{b} + \operatorname{div}^a (\mathbf{T}^\alpha)] d^{(\alpha)} V \quad (176)$$

Hence the equation of linear motion, expressing the balance of linear momentum, is defined by:

$$\operatorname{div}^a [\mathbf{T}^\alpha] + \rho \mathbf{b} - \rho \dot{\mathbf{v}} = 0 \quad (177)$$

It should be pointed out that div^a has already been defined by Eq. 96 and is different from the common definition of the divergence.

Following similar steps as in the conventional case, the balance of rotational momentum yields the symmetry of Cauchy stress tensor.

Balance of Rotational Momentum Principle

Following similar procedure as in the conventional case, we may end up to the symmetry property of the fractional stress tensor, i.e.:

$$\mathbf{T}^\alpha = (\mathbf{T}^\alpha)^T \quad (178)$$

Fractional Zener Viscoelastic Model

The Integer Viscoelastic Model

It was proposed by Zener consists of the elastic springs with elastic constraints E_1 and E_2 and the viscous dashpot having a viscosity constant C . Then the standard linear solid model called Zener viscoelastic model is indicated in Fig. 7.

The three-parameter Zener model (Fig. 7) yields the following constitutive equation:

$$\left[1 + \frac{C}{E_1 + E_2} \frac{d}{dt} \right] \sigma(t) = \frac{E_2}{E_1 + E_2} \left[E_1 + C \frac{d}{dt} \right] \varepsilon(t) \quad (179)$$

In this case the creep and relaxation functions take the form:

$$J(t) = \frac{E_1 + E_2}{E_1 E_2} \{ 1 - \exp[-(t/\tau_\varepsilon)] \} \quad (180)$$

$$G(t) = \frac{E_1 E_2}{E_1 + E_2} \{ 1 + \exp[-(t/\tau_\sigma)] \} \quad (181)$$

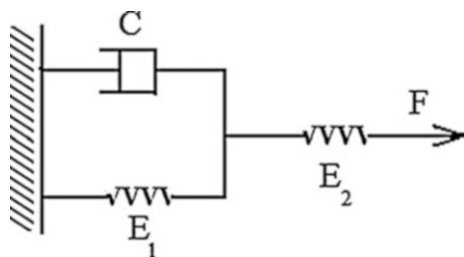
where

$$\tau_\varepsilon = \frac{C}{E_1}, \tau_\sigma = \frac{C}{E_1 + E_2}. \quad (182)$$

For the viscoelastic materials with $E_1 = 0.16 \cdot 10^9$ N/m, $E_2 = 1.5 \cdot 10^9$ N/m and $C = 16 \cdot 10^9$ Ns/m, and the initial condition $y_0 = 0$, the function of the compliance with respect to time is shown in Fig. 8.

Furthermore, the function of the relaxation modulus $G(t)$ with respect to time t is shown in Fig. 9:

Fig. 7 The three-parameter Zener model



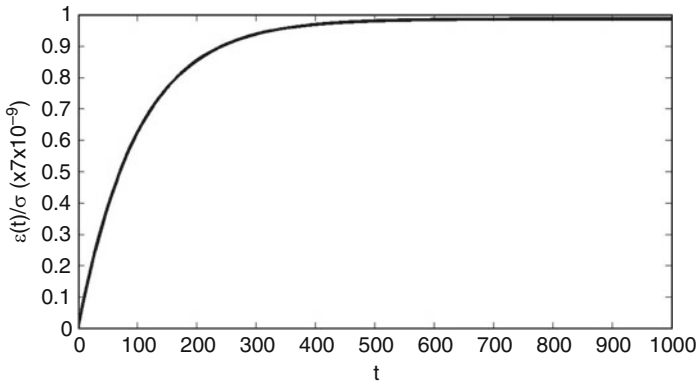


Fig. 8 The variation of the compliance with respect to time for the integer Zener viscoelastic model

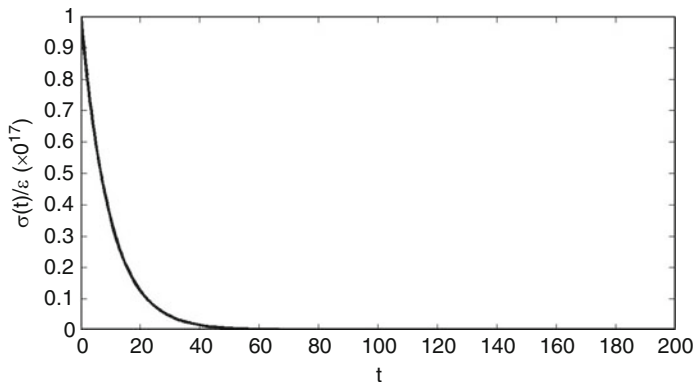


Fig. 9 Variation of the relaxation modulus with respect to time for the integer Zener viscoelastic model

The Fractional Order Derivative Viscoelastic Models

Viscoelastic models have been proposed using fractional time derivatives, just to take into account the material memory effects. The fractional Zener model introduced by Bagley and Torvik (1986) has been extensively studied in the literature (see Atanackovic 2002 and Sabatier et al. 2007). Consequently, the fractional Zener model has been proposed, using fractional time derivatives and more specifically the Caputo time derivatives. Therefore the constitutive equation for the model is:

$$\left[1 + \frac{C}{E_1 + E_2} \frac{d^\alpha}{dt^\alpha} \right] \sigma(t) = \frac{E_2}{E_1 + E_2} \left[E_1 + C \frac{d^\alpha}{dt^\alpha} \right] \varepsilon(t) \quad (183)$$

with creep compliance and relaxation modulus:

$$J(t) = \frac{1}{E_2} + \frac{1}{E_1} \{1 - E_\alpha [-(t/\tau_\epsilon)^\alpha]\} \tag{184}$$

$$G(t) = \frac{E_2}{E_1 + E_2} \{E_0 + E_1 E_\alpha [-(t/\tau_\epsilon)^\alpha]\} \tag{185}$$

where $\tau_\epsilon^\alpha = \frac{C}{E_1}, \tau_\sigma^\alpha = \frac{C}{E_1 + E_2}$.

For the viscoelastic materials with $E_1 = 0.16 \cdot 10^9$ N/m, $E_2 = 1.5 \cdot 10^9$ N/m, $C = 16 \cdot 10^9$ Ns/m, and the initial condition $y_0 = 0$, the function of the compliance with respect to time is shown in Fig. 10.

Furthermore, the function of the relaxation modulus $G(t)$ with respect to time t is shown in Fig. 11. It is shown that the lower the fractional order the slower the convergence to the final zero value of the relaxation.

Proposed Fractional Viscoelastic Zener Model

Since only Leibnitz L-fractional derivatives have physical sense, the fractional viscoelastic equations should be expressed in terms of L-fractional derivatives. Therefore the L-fractional Zener model should be expressed by:

$$\left[1 + \frac{C}{E_1 + E_2} {}^L_0 D_t^\alpha \right] \sigma(t) = \frac{E_2}{E_1 + E_2} [E_1 + C {}^L_0 D_t^\alpha] \epsilon(t) \tag{186}$$

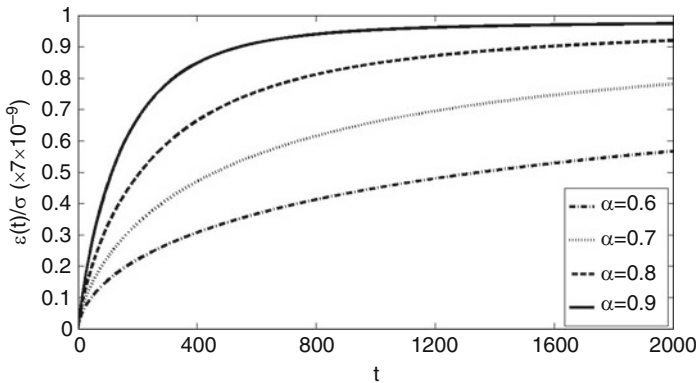


Fig. 10 The variation of the compliance with respect to time for various values of the fractional order for the existing fractional models

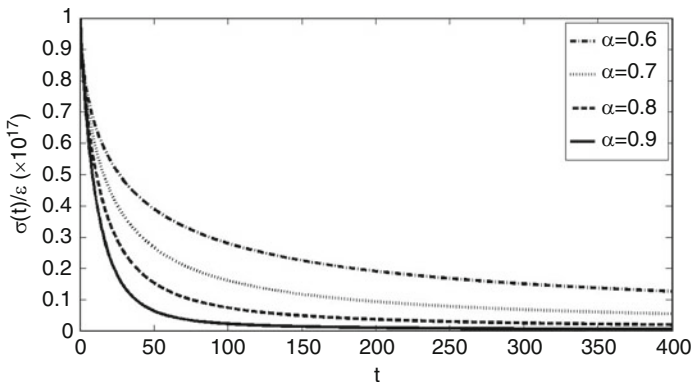


Fig. 11 The variation of the relaxation with respect to time for various values of the fractional order for the existing fractional models

Hence for constant applied stress the compliance

$$J(t) = \frac{\varepsilon(t)}{\sigma} \tag{187}$$

expressing the variation of strain with respect to time is given by the equation:

$$\left[E_1 + C_0^L D_t^\alpha \right] J(t) = \frac{E_1 + E_2}{E_2} \tag{188}$$

Therefore we have to solve the fractional linear equation:

$$C_0^L D_t^\alpha y(t) + E_1 y(t) = 1 + \frac{E_1}{E_2} \tag{189}$$

Looking now for a solution of the type:

$$y(t) = \sum_{k=0}^{\infty} y_k t^k \tag{190}$$

and substituting $y(t)$ from Eq. 190 to the governing compliance, Eq. 189 we get:

$$\sum_{k=0}^{\infty} C y_{k+1} \frac{\Gamma(2-\alpha) \Gamma(\kappa+2)}{\Gamma(\kappa+2-\alpha)} t^k + \sum_{k=0}^{\infty} E_1 y_k t^k = 1 + \frac{E_1}{E_2} \tag{191}$$

Since the algebraic Eq. 191 is valid for any t the various coefficients y_k are defined by:

$$y_1 = \frac{1}{C} \left(-E_1 y_0 + 1 + \frac{E_1}{E_2} \right) \tag{192}$$

$$y_{k+1} = -\frac{E_1}{C} \frac{\Gamma(\kappa + 2 - \alpha)}{\Gamma(\kappa + 2) \Gamma(2 - \alpha)} y_k, \forall k \geq 1 \tag{193}$$

Hence

$$y_k = \left(-\frac{E_1}{C \Gamma(2 - \alpha)} \right)^{k-1} \prod_{m=1}^{k-1} \frac{\Gamma(m + 2 - \alpha)}{\Gamma(m + 2)} y_1, \forall k \geq 2 \tag{194}$$

For the viscoelastic materials with $E_1 = 0.16 \cdot 10^9$ N/m, $E_2 = 1.5 \cdot 10^9$ N/m, $C = 16 \cdot 10^9$ Ns/m, and the initial condition $y_0 = 0$; the function of the compliance with respect to time is shown in Fig. 6.

It is clear from Fig. 6 that the fractional order has influence upon the time of convergence of the compliance modulus to the final value. The lower the value of the fractional order the slower the convergence of the compliance modulus to the final value.

Now, proceeding to the relaxation behavior of the Fractional Zener Viscoelastic model, we consider constant strain $\varepsilon(t) = \varepsilon$, then the governing Eq. 185 becomes:

$$C_0^L D_t^\alpha \left(\frac{\sigma(t)}{\varepsilon} \right) + (E_1 + E_2) \left(\frac{\sigma(t)}{\varepsilon} \right) = E_1 E_2 \tag{195}$$

For the relaxation modulus $y(t) = G(t) = \frac{\sigma(t)}{\varepsilon}$, the Eq. 197 above takes the form:

$$C_0^L D_t^\alpha y(t) + (E_1 + E_2) y(t) = E_1 E_2 \tag{196}$$

Looking for solution of the type

$$y(t) = \sum_{k=0}^{\infty} y_k t^k \tag{197}$$

and substituting in Eq. 196 we get:

$$\sum_{k=0}^{\infty} C y_{k+1} \frac{\Gamma(2 - \alpha) \Gamma(k + 2)}{\Gamma(k + 2 - \alpha)} t^k + \sum_{k=0}^{\infty} (E_1 + E_2) y_k t^k = E_1 E_2 \tag{198}$$

Since Eq. 198 is valid for any t, it is an identity. Hence:

$$y_1 = -\frac{E_1 + E_2}{C} y_0 + \frac{E_1 E_2}{C} \tag{199}$$

$$y_{k+1} = -\frac{E_1 + E_2}{C} \frac{\Gamma(m + 2 - \alpha)}{\Gamma(2 - \alpha) \Gamma(k + 2)} y_1, \forall k \geq 2 \tag{200}$$

Those relations yield:

$$y_k = \left(-\frac{E_1 + E_2}{c \Gamma(2 - \alpha)} \right)^{k-1} \prod_{m=1}^{k-1} \frac{\Gamma(m + 2 - \alpha)}{\Gamma(m + 2)} y_1, \forall k \geq 2. \tag{201}$$

For the viscoelastic materials with $E_1 = 0.16 \cdot 10^9$ N/m, $E_2 = 1.5 \cdot 10^9$ N/m, and $C = 16 \cdot 10^9$ Ns/m, the function of the relaxation modulus with respect to time is shown in Fig. 7.

It is clear from Fig. 7 that the fractional order has influence upon the time of convergence of the relaxation modulus to the zero value. The lower the value of the fractional order the slower the convergence of the relaxation modulus to the zero value.

Comparison of the Three Viscoelastic Models

For the viscoelastic materials with $E_1 = 0.16 \cdot 10^9$ N/m, $E_2 = 1.5 \cdot 10^9$ N/m, and $C = 16 \cdot 10^9$ Ns/m, and for two different values of the fractional orders $a = 0.5$, $a = 0.7$, and $a = 0.9$ the (Figs. 8, 9, and 10) show the behavior of the compliance modulus and (Figs. 11, 12, 13, 14, 15, and 16) the behavior of the relaxation modulus.

Comparing (Figs. 17, 18, and 19) the behavior of the proposed model exhibits a more mild convergence for the compliance and relaxation moduli to the final values as far as to Caputo’s viscoelastic models are concerned. Further, the lower

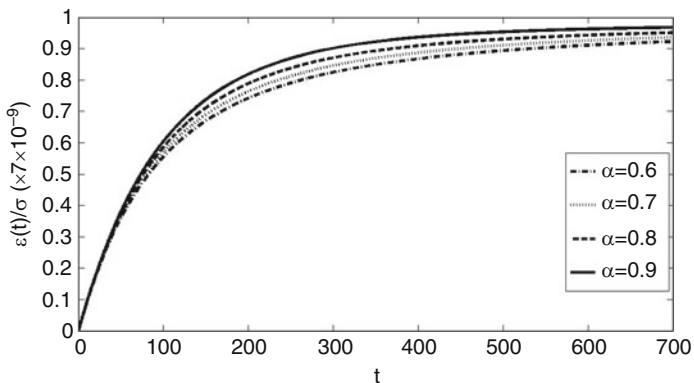


Fig. 12 The variation of the compliance with respect to time for various values of the fractional order for the existing fractional models

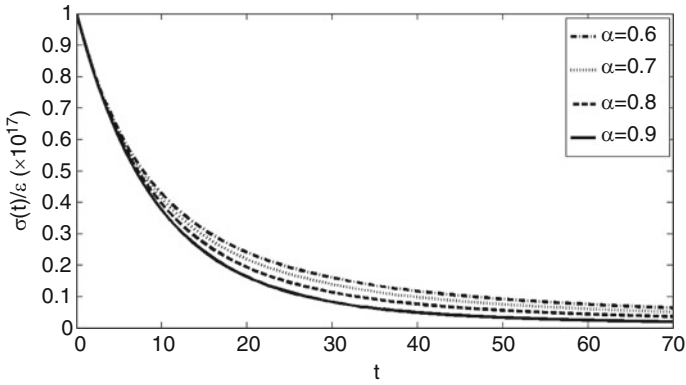


Fig. 13 Variation of the relaxation modulus with respect to time for various values of the fractional order for the existing fractional models

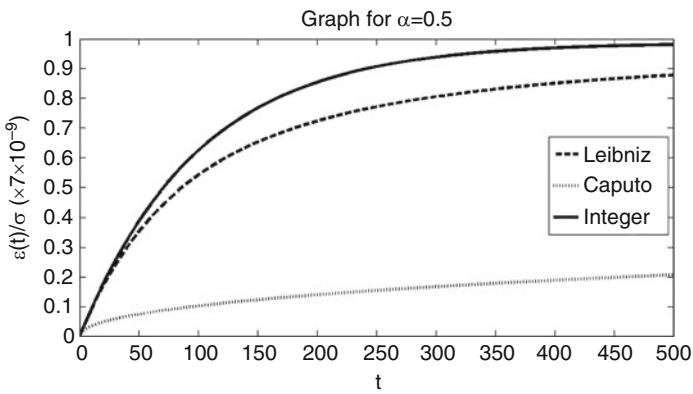


Fig. 14 The variation of the compliance with respect to time for fractional order $a = 0.5$

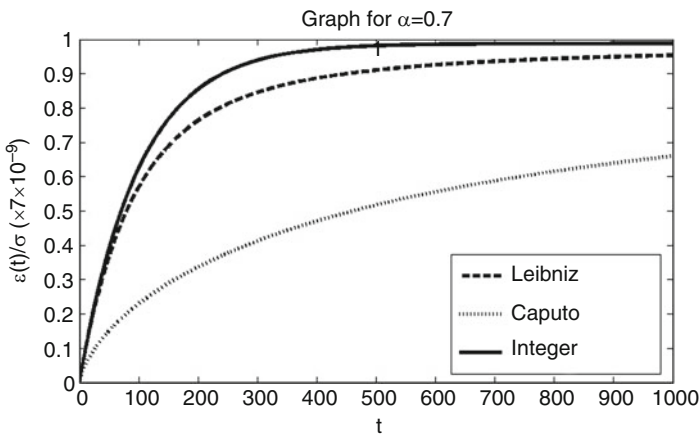


Fig. 15 The variation of the compliance with respect to time for fractional order $a = 0.7$

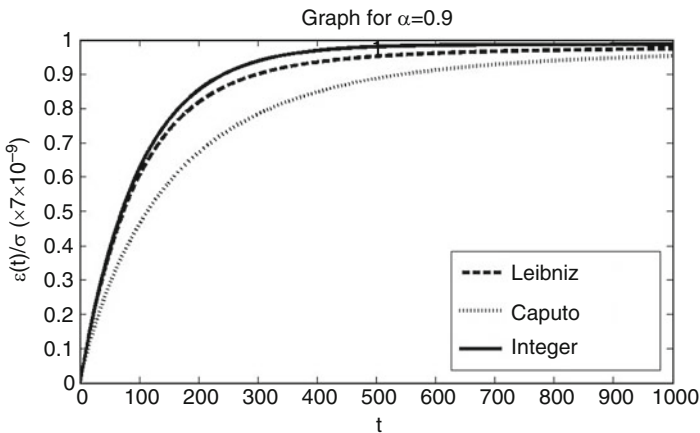


Fig. 16 The variation of the compliance with respect to time for fractional order $a = 0.9$

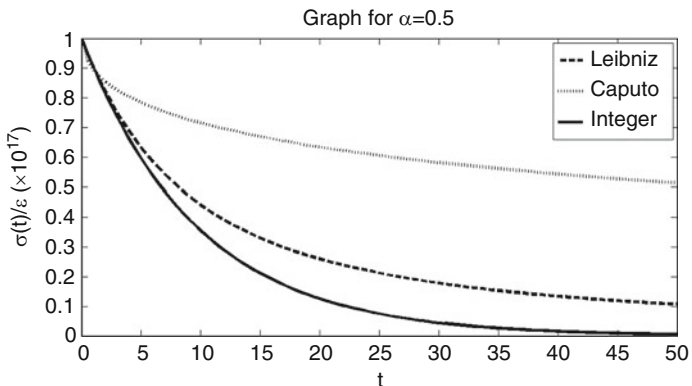


Fig. 17 Variation of the relaxation modulus with respect to time for fractional order $a = 0.5$

the fractional order of the model the slower the convergence to the final values for the same proposed model.

Conclusion: Further Research

Correcting the picture of fractional differential of a function, the fractional tangent space of a manifold was defined, introducing also Leibniz’s L-fractional derivative that is the only one having physical meaning. Further, the L-fractional chain rule is imposed, that is necessary for the existence of fractional differential. After establishing the fractional differential of a function, the theory of fractional differential geometry of curves is developed. In addition, the basic forms concerning the first and second differential forms of the surfaces were defined, through the

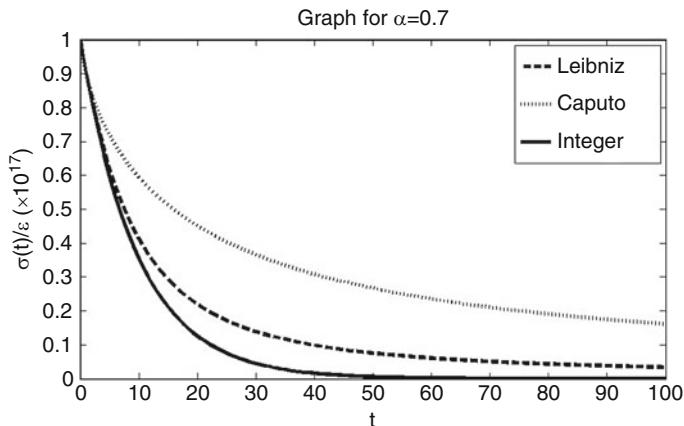


Fig. 18 Variation of the relaxation modulus with respect to time for fractional order $a = 0.7$

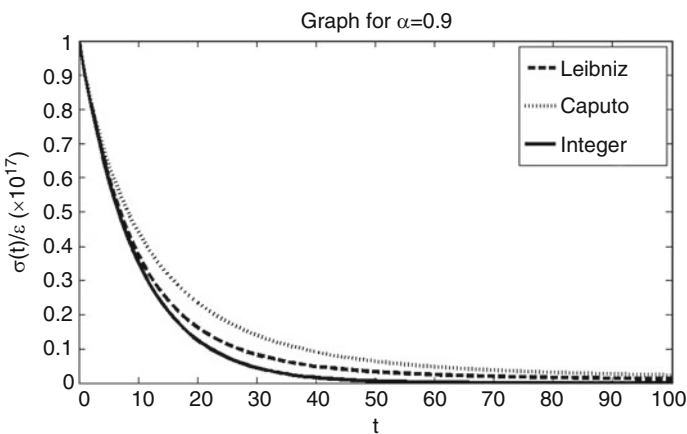


Fig. 19 Variation of the relaxation modulus with respect to time for fractional order $a = 0.9$

tangent spaces defined earlier, having mathematical meaning without any confusion, contrary to the existing procedures. Further the field theorems have been outlined in an accurate manner that may not cause confusion in their applications. The present work will help in discussing many applications concerning mechanics, quantum mechanics, and relativity, that need a clear description, based upon the fractional differential geometry. Moreover, the basic theorems of fractional vector calculus have been revised along with the basic concepts of fractional continuum mechanics. Especially the concept of fractional strain has been pointed out, because it is widely used in various places in a quite different way. The present analysis may be useful for solving updated problems in Mechanics and especially for lately proposed theories such as peridynamic theory. Finally, viscoelasticity models are studied using L-fractional derivatives. Those models exhibit milder behavior from the ones formulated through Caputo derivatives, compared to the conventional (integer)

models. In the present study, only the Zener viscoelastic model was discussed and its behavior concerning the compliance and relaxation moduli was studied.

References

- F.B. Adda, Interpretation geometrique de la differentiability et du gradient d'ordre reel. CR Acad. Sci. Paris **326**(Serie I), 931–934 (1998)
- F.B. Adda, The differentiability in the fractional calculus. Nonlinear Anal. **47**, 5423–5428 (2001)
- O.P. Agrawal, A general finite element formulation for fractional variational problems. J. Math. Anal. Appl. **337**, 1–12 (2008)
- E.C. Aifantis, Strain gradient interpretation of size effects. Int. J. Fract. **95**, 299–314 (1999)
- E.C. Aifantis, Update in a class of gradient theories. Mech. Mater. **35**, 259–280 (2003)
- E.C. Aifantis, On the gradient approach – relations to Eringen's nonlocal theory. Int. J. Eng. Sci. **49**, 1367–1377 (2011)
- H. Askes, E.C. Aifantis, Gradient elasticity in statics and dynamics: an overview of formulations, length scale identification procedures, finite element implementations and new results. Int. J. Solids Struct. **48**, 1962–1990 (2011)
- T.M. Atanackovic, A generalized model for the uniaxial isothermal deformation of a viscoelastic body. Acta Mech. **159**, 77–86 (2002)
- T.M. Atanackovic, B. Stankovic, Dynamics of a viscoelastic rod of fractional derivative type. ZAMM **82**(6), 377–386 (2002)
- T.M. Atanackovic, S. Konjik, S. Philipovic, Variational problems with fractional derivatives. Euler-Lagrange equations. J. Phys. A Math. Theor. **41**, 095201 (2008)
- R.L. Bagley, P.J. Torvik, Fractional calculus model of viscoelastic behavior. J. Rheol. **30**, 133–155 (1986)
- N. Bakhvalov, G. Panasenko, *Homogenisation: Averaging Processes in Periodic Media* (Kluwer, London, 1989)
- A. Balankin, B. Elizarratza, Hydrodynamics of fractal continuum flow. Phys. Rev. E **85**, 025302(R) (2012)
- D. Baleanu, K. Golmankhaneh Ali, K. Golmankhaneh Alir, M.C. Baleanu, Fractional electromagnetic equations using fractional forms. Int. J. Theor. Phys. **48**(11), 3114–3123 (2009)
- D.D. Baleanu, H. Srivastava, V. Daftardar-Gezzi, C. Li, J.A.T. Machado, Advanced topics in fractional dynamics. Adv. Mat. Phys. Article ID 723496 (2013)
- H. Beyer, S. Kempfle, Definition of physically consistent damping laws with fractional derivatives. ZAMM **75**(8), 623–635 (1995)
- G. Calcani, Geometry of fractional spaces. Adv. Theor. Math. Phys. **16**, 549–644 (2012)
- A. Carpinteri, B. Chiaia, P. Cornetti, Static-kinematic duality and the principle of virtual work in the mechanics of fractal media. Comput. Methods Appl. Mech. Eng. **191**, 3–19 (2001)
- A. Carpinteri, P. Cornetti, A. Sapora, Static-kinematic fractional operator for fractal and non-local solids. ZAMM **89**(3), 207–217 (2009)
- A. Carpinteri, P. Cornetti, A. Sapora, A fractional calculus approach to non-local elasticity. Eur. Phys. J. Spec Top **193**, 193–204 (2011)
- D.R.J. Chillingworth, *Differential Topology with a View to Applications* (Pitman, London, 1976)
- M. Di Paola, G. Failla, M. Zingales, Physically-based approach to the mechanics of strong non-local linear elasticity theory. J. Elast. **97**(2), 103–130 (2009)
- C.S. Drapaca, S. Sivaloganathan, A fractional model of continuum mechanics. J. Elast. **107**, 107–123 (2012)
- A.C. Eringen, *Nonlocal Continuum Field Theories* (Springer, New York, 2002)
- G.I. Evangelatos, *Non Local Mechanics in the Time and Space Domain-Fracture Propagation via a Peridynamics Formulation: A Stochastic/Deterministic Perspective*. Thesis, Houston Texas, 2011

- G.I. Evangelatos, P.D. Spanos, Estimating the ‘In-Service’ modulus of elasticity and length of polyester mooring lines via a non linear viscoelastic model governed by fractional derivatives. *ASME 2012 Int. Mech. Eng. Congr. Expo.* **8**, 687–698 (2012)
- J. Feder, *Fractals* (Plenum Press, New York, 1988)
- A.K. Goldmankhaneh, A.K. Goldmankhaneh, D. Baleanu, Lagrangian and Hamiltonian mechanics. *Int. J. Theor. Rhys.* **52**, 4210–4217 (2013)
- K. Golmankhaneh Ali, K. Golmankhaneh Alir, D. Baleanu, About Schrodinger equation on fractals curves imbedding in R^3 . *Int. J. Theor. Phys.* **54**(4), 1275–1282 (2015)
- H. Guggenheimer, *Differential Geometry* (Dover, New York, 1977)
- G. Jumarie, An approach to differential geometry of fractional order via modified Riemann-Liouville derivative. *Acta Math. Sin. Engl. Ser.* **28**(9), 1741–1768 (2012)
- A.A. Kilbas, H.M. Srivastava, J.J. Trujillo, *Theory and Applications of Fractional Differential Equations* (Elsevier, Amsterdam, 2006)
- K.A. Lazopoulos, On the gradient strain elasticity theory of plates. *Eur. J. Mech. A/Solids* **23**, 843–852 (2004)
- K.A. Lazopoulos, Nonlocal continuum mechanics and fractional calculus. *Mech. Res. Commun.* **33**, 753–757 (2006)
- K.A. Lazopoulos, in *Fractional Vector Calculus and Fractional Continuum Mechanics, Conference “Mechanics though Mathematical Modelling”, celebrating the 70th birthday of Prof. T. Atanackovic*, Novi Sad, 6–11 Sept, Abstract, p. 40 (2015)
- A.K. Lazopoulos, On fractional peridynamic deformations. *Arch. Appl. Mech.* **86**(12), 1987–1994 (2016a)
- K.A. Lazopoulos, in *Fractional Differential Geometry of Curves and Surfaces, International Conference on Fractional Differentiation and Its Applications (ICFDA 2016)*, Novi Sad (2016b)
- A.K. Lazopoulos, *On Fractional Peridynamic Deformations, International Conference on Fractional Differentiation and Its Applications, Proceedings ICFDA 2016*, Novi Sad (2016c)
- K.A. Lazopoulos, A.K. Lazopoulos, Bending and buckling of strain gradient elastic beams. *Eur. J. Mech. A/Solids* **29**(5), 837–843 (2010)
- K.A. Lazopoulos, A.K. Lazopoulos, On fractional bending of beams. *Arch. Appl. Mech.* (2015). <https://doi.org/10.1007/S00419-015-1083-7>
- K.A. Lazopoulos, A.K. Lazopoulos, Fractional vector calculus and fractional continuum mechanics. *Prog. Fract. Diff. Appl.* **2**(1), 67–86 (2016a)
- K.A. Lazopoulos, A.K. Lazopoulos, On the fractional differential geometry of curves and surfaces. *Prog. Fract. Diff. Appl.*, No **2**(3), 169–186 (2016b)
- G.W. Leibnitz, Letter to G. A. L’Hospital. *Leibnizen Mathematische Schriften.* **2**, 301–302 (1849)
- Y. Liang, W. Su, Connection between the order of fractional calculus and fractional dimensions of a type of fractal functions. *Anal. Theory Appl.* **23**(4), 354–362 (2007)
- J. Liouville, Sur le calcul des differentielles a indices quelconques. *J. Ec. Polytech.* **13**, 71–162 (1832)
- H.-S. Ma, J.H. Prevost, G.W. Sherer, Elasticity of dlca model gels with loops. *Int. J. Solids Struct.* **39**, 4605–4616 (2002)
- F. Mainardi, *Fractional Calculus and Waves in Linear Viscoelasticity* (Imperial College Press, London, 2010)
- M. Meerschaert, J. Mortensen, S. Wheatcraft, Fractional vector calculus for fractional advection–dispersion. *Physica A* **367**, 181–190 (2006)
- R.D. Mindlin, Second gradient of strain and surface tension in linear elasticity. *Int. Jnl. Solids & Struct.* **1**, 417–438 (1965)
- W. Noll, A mathematical theory of the mechanical behavior of continuous media. *Arch. Rational Mech. Anal.* **2**, 197–226 (1958/1959)
- K.B. Oldham, J. Spanier, *The Fractional Calculus* (Academic, New York, 1974)
- I. Podlubny, *Fractional Differential Equations (An Introduction to Fractional Derivatives Fractional Differential Equations, Some Methods of Their Solution and Some of Their Applications)* (Academic, San Diego, 1999)
- I.R. Porteous, *Geometric Differentiation* (Cambridge University Press, Cambridge, 1994)

- B. Riemann, Versuch einer allgemeinen Auffassung der Integration und Differentiation, in *Gesammelte Werke*, vol. 62 (1876)
- F. Riewe, Nonconservative Lagrangian and Hamiltonian mechanics. *Phys. Rev. E* **53**(2), 1890–1899 (1996)
- F. Riewe, Mechanics with fractional derivatives. *Phys. Rev. E* **55**(3), 3581–3592 (1997)
- J. Sabatier, O.P. Agrawal, J.A. Machado, *Advances in Fractional Calculus (Theoretical Developments and Applications in Physics and Engineering)* (Springer, The Netherlands, 2007)
- S.G. Samko, A.A. Kilbas, O.I. Marichev, *Fractional Integrals and Derivatives: Theory and Applications* (Gordon and Breach, Amsterdam, 1993)
- S.A. Silling, Reformulation of elasticity theory for discontinuities and long-range forces. *J. Mech. Phys. Solids* **48**, 175–209 (2000)
- S.A. Silling, R.B. Lehoucq, Peridynamic theory of solid mechanics. *Adv. Appl. Mech.* **44**, 175–209 (2000)
- S.A. Silling, M. Zimmermann, R. Abeyaratne, Deformation of a peridynamic bar. *J. Elast.* **73**, 173–190 (2003)
- W. Sumelka, Non-local Kirchhoff-Love plates in terms of fractional calculus. *Arch. Civil and Mech. Eng.* **208** (2014). <https://doi.org/10.1016/j.acme2014.03.006>
- V.E. Tarasov, Fractional vector calculus and fractional Maxwell's equations. *Ann. Phys.* **323**, 2756–2778 (2008)
- V.E. Tarasov, *Fractional Dynamics: Applications of Fractional Calculus to Dynamics of Particles, Fields and Media* (Springer, Berlin, 2010)
- R.A. Toupin, Theories of elasticity with couple stress. *Arch. Ration. Mech. Anal.* **17**, 85–112 (1965)
- C. Truesdell, *A First Course in Rational Continuum Mechanics*, vol 1 (Academic, New York, 1977)
- C. Truesdell, W. Noll, The non-linear field theories of mechanics, in *Handbuch der Physik*, vol. III/3, ed. by S. Flügge (Springer, Berlin, 1965)
- I. Vardoulakis, G. Exadaktylos, S.K. Kourkoulis, Bending of a marble with intrinsic length scales: a gradient theory with surface energy and size effects. *J. Phys. IV* **8**, 399–406 (1998)
- H.M. Wyss, A.M. Deliormanli, E. Tervoort, L.J. Gauckler, Influence of microstructure on the rheological behaviour of dense particle gels. *AIChE J.* **51**, 134–141 (2005)
- K. Yao, W.Y. Su, S.P. Zhou, On the connection between the order of fractional calculus and the dimensions of a fractal function. *Chaos, Solitons Fractals* **23**, 621–629 (2005)



Continuum Homogenization of Fractal Media

25

Martin Ostoja-Starzewski, Jun Li, and Paul N. Demmie

Contents

Introduction	906
Homogenization of Fractal Media	907
Mass Power Law and Product Measure	907
Product Measure	909
Fractional Integral Theorems and Fractal Derivatives	911
Vector Calculus on Anisotropic Fractals	913
Homogenization Process for Fractal Media	914
Continuum Mechanics of Fractal Media	916
Fractal Continuity Equation	917
Fractal Linear Momentum Equation	917
Fractal Angular Momentum Equation	918
Fractal Energy Equation	920
Fractal Second Law of Thermodynamics	921
Fractal Wave Equations	923
Elastodynamics of a Fractal Timoshenko Beam	923
Elastodynamics in 3D	925
Related Topics	926

Sandia National Laboratories is a multimission laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.

M. Ostoja-Starzewski (✉)

Department of Mechanical Science and Engineering, Institute for Condensed Matter Theory and Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, IL, USA
e-mail: martinos@illinois.edu

J. Li

Department of Mechanical Engineering, University of Massachusetts, Dartmouth, MA, USA

P. N. Demmie

Sandia National Laboratories, Albuquerque, NM, USA

Extremum and Variational Principles in Fractal Bodies	926
Fracture in Elastic-Brittle Fractal Solids	928
Closure	933
References	933

Abstract

This chapter reviews the modeling of fractal materials by homogenized continuum mechanics using calculus in non-integer dimensional spaces. The approach relies on expressing the global balance laws in terms of fractional integrals and, then, converting them to integer-order integrals in conventional (Euclidean) space. Via localization, this allows development of local balance laws of fractal media (continuity, linear and angular momenta, energy, and second law) and, in case of elastic responses, formulation of wave equations in several settings (1D and 3D wave motions, fractal Timoshenko beam, and elastodynamics under finite strains). Next, follows an account of extremum and variational principles, and fracture mechanics. In all the cases, the derived equations for fractal media depend explicitly on fractal dimensions and reduce to conventional forms for continuous media with Euclidean geometries upon setting the dimensions to integers.

Keywords

Balance laws · Fractal · Fractional calculus · Fractal derivative · Homogenization

Introduction

It has been observed by Benoît Mandelbrot (1982) that many natural objects are statistically self-similar and “broken” in space (or time) and exhibit non-smooth or highly irregular features. Mandelbrot called such objects *fractals*. Examples include coastlines, porous media, cracks, turbulent flows, clouds, mountains, lightning bolts, snowflakes, melting ice, and even parts of living entities such as the neural structure or the surface of the human brain (Barnsley 1993; Le Méhauté 1991; Hastings and Sugihara 1993; Falconer 2003; Tarasov 2005a). There are also fractals in time – signals, processes, and musical compositions – but we shall not concern ourselves with them here.

Mathematical fractal sets are characterized by a Hausdorff dimension D , which is the scaling exponent characterizing the fractal pattern’s power law. Physical fractals can be modeled only for some finite range of length scales within the lower and upper cutoffs by mathematical ones. These objects are called *pre-fractals*. While the mechanics of fractal and pre-fractal media is still in an early developing stage, the new field of *fractal mechanics* can already generate elegant models (Tarasov 2005b; Ostoja-Starzewski 2007, 2008a, 2009; Joumaa and Ostoja-Starzewski 2011; Li and Ostoja-Starzewski 2009a).

The approach we employ is based on calculus in non-integer dimensional spaces (i.e., spaces embedded in Euclidean space), thus allowing a homogenization; by a slight abuse of terminology stemming from quantum mechanics (Li and Ostoja-Starzewski 2009b), this method has originally been called “dimensional regularization.” With this approach the fractional integrals over fractal sets (i.e., fractal material domains) are transformed to equivalent continuous integrals over Euclidean sets (Jumarie 2009). This transformation produces balance laws that are expressed in continuous form, thereby simplifying their mathematical manipulation both analytically and computationally. A product measure is used to achieve this transformation.

The idea goes back to V.E. Tarasov (2005a,b,c), who developed continuum-type equations for conservation of mass, linear and angular momentum, and energy of fractals and studied several fluid mechanics and wave problems (Tarasov 2005a,b,c, 2010). An advantage of this approach is that it admits upper and lower cutoffs of fractal scaling, so that one effectively deals with a physical pre-fractal rather than a purely mathematical fractal lacking any cutoffs. The original formulation of Tarasov was based on the Riesz measure, which is more appropriate for isotropic fractal media than for anisotropic fractal media. To represent more general heterogeneous media, Li and Ostoja-Starzewski introduced a model based on a product measure (Li and Ostoja-Starzewski 2010, 2011; Ostoja-Starzewski et al. 2016). Since this measure has different fractal dimensions in different directions, it grasps the anisotropy of fractal geometry better than the Tarasov formulation for a range of length scales between the lower and upper cutoffs (Li and Ostoja-Starzewski 2010; Ostoja-Starzewski et al. 2016). The great promise is that the conventional requirement of continuum mechanics, the separation of scales, can be removed with continuum-type field equations still employed. This approach was applied, among others, to thermomechanics with internal variables, extremum principles of elasticity and plasticity, turbulence in fractal porous media, dynamics of fractal beams, fracture mechanics, and thermoelasticity (Ostoj-Starzewski 2007, 2008a, 2009; Jumarie 2009; Oldham and Spanier 1974; Ostoja-Starzewski et al. 2014).

Homogenization of Fractal Media

Mass Power Law and Product Measure

The basic approach to homogenization of fractal media by continua originated with Tarasov (2005a,b, 2010). He started to work in the setting where the mass obeys a power law

$$m(R) \sim R^D, \quad D < 3, \quad (1)$$

with R being the length scale of measurement (or resolution) and D the fractal dimension of mass in the three-dimensional (3D) Euclidean space \mathbb{E}^3 . Note that the relation (1) can be applied to a *pre-fractal*, i.e., a fractal-type, physical object with lower and upper cutoffs. More specifically, Tarasov used a fractional integral to represent mass in a region \mathcal{W} embedded in \mathbb{E}^3 . In the subsequent work, we focused on a general anisotropic, fractal medium governed, in place of (1), by a more general power law relation with respect to each coordinate (Li and Ostoja-Starzewski 2010, 2011; Ostoja-Starzewski et al. 2016) (which, in fact, had originally been recognized by Tarasov)

$$m(x_1, x_2, x_3) \sim x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3}. \quad (2)$$

Then, the mass is specified via a *product measure*

$$m(\mathcal{W}) = \int_{\mathcal{W}} \rho(x_1, x_2, x_3) dl_{\alpha_1}(x_1) dl_{\alpha_2}(x_2) dl_{\alpha_3}(x_3), \quad (3)$$

while the length measure along each coordinate is given through transformation coefficients $c_1^{(k)}$

$$dl_{\alpha_k}(x_k) = c_1^{(k)}(\alpha_k, x_k) dx_k, \quad k = 1, 2, 3 \quad (\text{no sum}). \quad (4)$$

Equation (3) implies that the mass fractal dimension D equals $\alpha_1 + \alpha_2 + \alpha_3$ along the diagonals, $|x_1| = |x_2| = |x_3|$, where each α_k plays the role of a fractal dimension in the direction x_k . While it is noted that, in other directions, the anisotropic fractal body's fractal dimension is not necessarily the sum of projected fractal dimensions, an observation from an established text on mathematics of fractals is recalled here (Falconer 2003): *Many fractals encountered in practice are not actually products, but are product-like*. In what follows, we expect the equality between $D = \alpha_1 + \alpha_2 + \alpha_3$ to hold for fractals encountered in practice, whereas a rigorous proof of this property remains an open research topic.

This formulation has four advantages over the original approach (Li and Ostoja-Starzewski 2009b, 2011):

1. It does not involve a left-sided fractional derivative (Riemann-Liouville), which does not give zero when applied to a constant function, a rather unphysical property.
2. The mechanics-type derivation of wave equations (in 3D, 2D, or 1D) yields the same result as that obtained from the variational-type derivation.
3. The 3D wave equation cleanly reduces to the 1D wave equation.
4. It offers a consistent way of handling not only generally anisotropic fractals but also isotropic ones, overall developing a systematic formulation of continuum mechanics. In particular, see the “fractal derivative” (13) below.

Product Measure

The relation (4) implies that the infinitesimal fractal volume element, dV_D , is

$$dV_D = dl_{\alpha_1}(x_1)dl_{\alpha_2}(x_2)dl_{\alpha_3}(x_3) = c_1^{(1)}c_1^{(2)}c_1^{(3)}dx_1dx_2dx_3 = c_3dV_3, \tag{5}$$

$$\text{with } c_3 = c_1^{(1)}c_1^{(2)}c_1^{(3)}.$$

In fact, it plays the role of a fractal representative volume element (RVE), which is mapped into the RVE of non-fractal (conventional) type. Note that (i) this map is different from the scaling of a random microstructure toward a deterministic volume element in non-fractal media (Ostoja-Starzewski et al. 2016), and (ii) the material spatial randomness is not explicitly introduced into the formulation discussed here.

For the surface transformation coefficient $c_2^{(k)}$, we consider a cubic volume element, $dV_3 = dx_1dx_2dx_3$, whose surface elements are specified by the normal vector along the axes i, j , or k in Fig. 1. Therefore, $c_2^{(k)}$ associated with the surface $S_d^{(k)}$ is

$$c_2^{(k)} = c_1^{(i)}c_1^{(j)} = c_3/c_1^{(k)}, \quad i \neq j, \quad i, j \neq k. \tag{6}$$

The sum $d^{(k)} = \alpha_i + \alpha_j, i \neq j, i, j \neq k$, is the fractal dimension of the surface $S_d^{(k)}$ along the diagonals $|x_i| = |x_j|$ in $S_d^{(k)}$. This equality is not necessarily true elsewhere, but is expected to hold for fractals encountered in practice (Falconer 2003) as discussed previously for the relationship between D and $\alpha_1 + \alpha_2 + \alpha_3$. Figure 1 illustrates the relationship among the line transformation coefficients $c_1^{(k)}$ and respective surface ($c_2^{(k)}$) and volume (c_3) transformation coefficients. We note that, when $D \rightarrow 3$, with each $\alpha_i \rightarrow 1$, the conventional concept of mass is recovered (Ostoja-Starzewski et al. 2014).

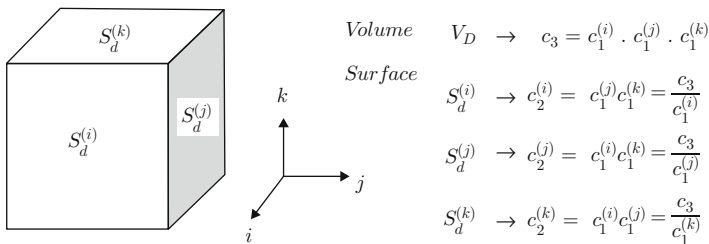


Fig. 1 Roles of the transformation coefficients $c_1^{(i)}, c_2^{(k)}$, and c_3 in homogenizing a fractal body of volume dV_D , surface dS_d , and lengths dl_{α} into a Euclidean parallelepiped of volume dV_3 , surface dS_2 , and lengths dx_k

We adopt the modified Riemann-Liouville fractional integral of Jumarie (2009, 2005) for $c_1^{(k)}$:

$$c_1^{(k)} = \alpha_k \left(\frac{l_k - x_k}{l_{k0}} \right)^{\alpha_k - 1}, \quad k = 1, 2, 3, \quad (\text{no sum on } k), \quad (7)$$

where l_k is the total length (integral interval) along x_k , while l_{k0} is the characteristic length in the given direction (e.g., the mean pore size). In the product measure formulation, the resolution length scale is

$$R = \sqrt{l_k l_k}. \quad (8)$$

Examining $c_1^{(k)}$ in two special cases, we observe:

1. Uniform mass: The mass is distributed isotropically in a cubic region with a power law relation (9). Denoting the reference mass density by ρ_0 and the cubic length by l , we obtain

$$m(\mathcal{W}) = \rho_0 l^{\alpha_1} l^{\alpha_2} l^{\alpha_3} / l_0^{D-3} = \rho_0 l^{\alpha_1 + \alpha_2 + \alpha_3} / l_0^{D-3} = \rho_0 l^D / l_0^{D-3}, \quad (9)$$

which is consistent with the mass power law (1). In general, however, $D \neq \alpha_1 + \alpha_2 + \alpha_3$.

2. Point mass: The distribution of mass is concentrated at one point, so that the mass density is denoted by the Dirac function $\rho(x_1, x_2, x_3) = m_0 \delta(x_1) \delta(x_2) \delta(x_3)$. The fractional integral representing mass becomes

$$m(\mathcal{W}) = \alpha_1 \alpha_2 \alpha_3 \frac{l^{\alpha_1 - 1} l^{\alpha_2 - 1} l^{\alpha_3 - 1}}{l_0^{D-3}} m_0 = \alpha_1 \alpha_2 \alpha_3 \left(\frac{l}{l_0} \right)^{D-3} m_0, \quad (10)$$

When $D \rightarrow 3$ ($\alpha_1, \alpha_2, \alpha_3 \rightarrow 1$), $m(\mathcal{W}) \rightarrow m_0$ and the conventional concept of point mass is recovered (Ostoja-Starzewski et al. 2014). Note that using the Riesz fractional integral is not well defined except when $D = 3$ (by letting $0^0 = 1$ in $m(\mathcal{W}) = \alpha_1 \alpha_2 \alpha_3 0^{D-3} m_0$), which, on the other hand, shows a non-smooth transition of mass with respect to its fractal dimension. This fact also supports our choice of the non-Riesz-type expressions for $c_1^{(k)}$ in (14).

Note that our expression for $c_1^{(k)}$ shows that the length dimension, and hence the mass m would involve an unusual physical dimension if it were replaced by $c_1^{(k)} = \alpha_k (l_k - x_k)^{\alpha_k - 1}$. This behavior is understandable since, mathematically, a fractal curve only exhibits a finite measure with respect to a fractal dimensional length unit (Mandelbrot 1982). Of course, in practice, we prefer physical quantities to have usual dimensions, and so we work with nondimensionalized coefficients $c_1^{(k)}$.

A simple generic example of an anisotropic, product-like fractal is the so-called *Carpinteri column* (Carpinteri et al. 2004), a parallelepiped domain in \mathbb{E}^3 , having

mathematically well-defined Hausdorff dimensions in all three directions. It has been proposed as a model of concrete columns which are essentially composite structures featuring oriented fractal-type microstructures. The square cross section of the column is a *Sierpiński carpet*, which is “fractally” swept along the longitudinal direction in conjunction with a *Cantor ternary set* (see Ostoja-Starzewski et al. 2014 for a discussion).

Another example is offered by the rings of Saturn. If considered as random fields, the rings possess (i) statistical stationarity in time, (ii) statistical isotropy in space, and (iii) statistical spatial nonstationarity. The reason for (i) is an extremely slow decay of rings relative to the time scale of orbiting around Saturn. The reason for (ii) is the obviously circular, albeit radially disordered and fractal pattern of rings; fractality is present in the radial but not polar coordinate (Li and Ostoja-Starzewski 2015). The reason for (iii) is the lack of invariance with respect to arbitrary shifts in Cartesian space which, on the contrary, holds true in, say, a basic model of turbulent velocity fields.

Fractional Integral Theorems and Fractal Derivatives

In order to develop continuum mechanics of fractal media, we introduce the notion of fractal derivatives with respect to the coordinate x_k and time t . The definitions of these derivatives follow naturally from the fractional generalization of two basic integral theorems that are employed in continuum mechanics (Tarasov 2005c, 2010; Li and Ostoja-Starzewski 2009b; Demmie and Ostoja-Starzewski 2011): Gauss theorem, which relates a volume integral to the surface integral over its bounding surface, and the Reynolds transport theorem, which provides an expression for the time rate of change of any volume integral in a continuous medium.

Consider the surface integral

$$\int_{\partial\mathcal{W}} \mathbf{f} \cdot \mathbf{n} dS_d = \int_{\partial\mathcal{W}} f_k n_k dS_d, \tag{11}$$

where $\mathbf{f}(= f_k)$ is any vector field and $\mathbf{n}(= n_k)$ is the outward normal vector field to the surface $\partial\mathcal{W}$ which is the boundary surface for some volume \mathcal{W} , and dS_d is the surface element in fractal space. The notation $(= A_k)$ is used to indicate that the A_k are components of the vector \mathbf{A} .

To compute (11), we relate the integral element $\mathbf{n} dS_d$ to its conventional surface element $\mathbf{n} dS_2$ in \mathbb{E}^3 via the fractal surface coefficients $c_2^{(k)}$, $k = 1, 2, 3$, as shown in Fig. 1. This figure shows that the infinitesimal element $\mathbf{n} dS_d$ can be expressed as a linear combination of the $n_k c_2^{(k)} dS_2$, $k = 1, 2, 3$ (no sum).

By the conventional Gauss theorem, and noting that $c_2^{(k)}$ does not depend on the coordinate x_k , (11) becomes

$$\int_{\partial\mathcal{W}} \mathbf{f} \cdot \mathbf{n} dS_d = \int_{\partial\mathcal{W}} f_k n_k dS_d = \int_{\mathcal{W}} [f_k c_2^{(k)}]_{,k} dV_3 = \int_{\mathcal{W}} \frac{f_{k,k}}{c_1^{(k)}} dV_D. \tag{12}$$

In (12) and elsewhere, we employ the usual convention that $(\cdot)_{,k}$ is the partial derivative of (\cdot) with respect to x_k . Next, based on (11) and the above, we define the *fractal derivative (fractal gradient)*, ∇_k^D as

$$\nabla^D \phi = \mathbf{e}_k \nabla_k^D \phi \quad \text{or} \quad \nabla_k^D \phi = \frac{1}{c_1^{(k)}} \frac{\partial \phi}{\partial x_k} \quad (\text{no sum on } k), \tag{13}$$

where \mathbf{e}_k is the base vector. With this definition, the *Gauss theorem for fractal media* becomes

$$\int_{\partial \mathcal{W}} \mathbf{f} \cdot \mathbf{n} dS_d = \int_{\mathcal{W}} \nabla_k^D f_k dV_D = \int_{\mathcal{W}} (\nabla^D \cdot \mathbf{f}) dV_D. \tag{14}$$

It is straightforward to show that the fractal operator, ∇_k^D , commutes with the fractional integral operator, is its inverse, and satisfies the product rule for differentiation (the Leibnitz property). Furthermore, the fractal derivative of a constant is zero. This latter property shows that a fractal derivative and a fractional derivative are not the same since the fractional derivative of a constant does not always equal to zero (Oldham and Spanier 1974).

To define the fractal material time derivative, we consider the fractional generalization of Reynolds transport theorem. Consider any quantity, P , accompanied by a moving fractal material system, \mathcal{W}_t , with velocity vector field \mathbf{v} ($= v_k$). The time derivative of the volume integral of P over \mathcal{W}_t is

$$\frac{d}{dt} \int_{\mathcal{W}_t} P dV_D. \tag{15}$$

Using the Jacobian (J) of the transformation between the current configuration (x_k) and the reference configuration (X_k), the relationship between the corresponding volume elements ($dV_D = J dV_D^0$), and the expression for the time derivative of J , it is straightforward to show that

$$\begin{aligned} \frac{d}{dt} \int_{\mathcal{W}_t} P dV_D &= \int_{\mathcal{W}_t} \left[\frac{\partial P}{\partial t} + (v_k P)_{,k} \right] dV_D \\ &= \int_{\mathcal{W}_t} \left[\frac{\partial P}{\partial t} + c_1^{(k)} \nabla_k^D (v_k P) \right] dV_D. \end{aligned} \tag{16}$$

The result given by the first equality is identical to the conventional representation. Hence, the fractal material time derivative and the conventional material time derivative are the same,

$$\left(\frac{d}{dt} \right)_D P = \frac{\partial P}{\partial t} + v_k P_{,k} = \frac{\partial P}{\partial t} + c_1^{(k)} v_k \nabla_k^D P. \tag{17}$$

Equation (17) is the *Reynolds theorem for fractal media*; its form is similar to the conventional one; an alternative form (Li and Ostoja-Starzewski 2010) of the fractional Reynolds transport theorem that involves surface integrals is different from the conventional one and rather complicated. This difference results from the fractal volume coefficient c_3 depending on all the coordinates, whereas in the derivation of the Gauss theorem (14), $c_2^{(k)}$ is independent of x_k .

Vector Calculus on Anisotropic Fractals

Motivated by the fractal derivative (13), we have the *fractal divergence* of a vector field (\mathbf{f})

$$\text{divf} = \nabla^D \cdot \mathbf{f} \quad \text{or} \quad \nabla_k^D f_k = \frac{1}{c_1^{(k)}} \frac{\partial f_k}{\partial x_k} \tag{18}$$

and the *fractal curl* operator

$$\text{curlf} = \nabla^D \times \mathbf{f} \quad \text{or} \quad e_{jki} \nabla_k^D f_i = e_{jki} \frac{1}{c_1^{(k)}} \frac{\partial f_i}{\partial x_k}. \tag{19}$$

The four fundamental identities of the conventional vector calculus can now be shown to carry over in terms of these new operators:

- (i) The divergence of the curl of a vector field \mathbf{f} :

$$\text{div} \cdot \text{curlf} = \mathbf{e}_m \nabla_m^D \cdot \mathbf{e}_j e_{jki} \nabla_k^D f_i = \frac{1}{c_1^{(j)}} \frac{\partial f}{\partial x_j} \left[e_{jki} \frac{1}{c_1^{(k)}} \frac{\partial f_i}{\partial x_k} \right] = 0. \tag{20}$$

where e_{ijk} is the permutation tensor.

- (ii) The curl of the gradient of a scalar field ϕ :

$$\text{curl} \times (\text{grad}\phi) = \mathbf{e}_i e_{ijk} \nabla_j^D (\nabla_k^D \phi) = \mathbf{e}_i e_{ijk} \frac{1}{c_1^{(j)}} \frac{\partial}{\partial x_j} \left[\frac{1}{c_1^{(k)}} \frac{\partial \phi}{\partial x_k} \right] = 0. \tag{21}$$

In both cases above, we can pull $1/c_1^{(k)}$ in front of the gradient because the coefficient $c_1^{(k)}$ is independent of x_j .

- (iii) The divergence of the gradient of a scalar field ϕ is written in terms of the fractal gradient as

$$\text{div} \cdot (\text{grad}\phi) = \nabla_j^D \cdot \nabla_k^D \phi = \frac{1}{c_1^{(j)}} \frac{\partial}{\partial x_j} \left[\frac{1}{c_1^{(j)}} \frac{\partial \phi}{\partial x_j} \right] = \frac{1}{c_1^{(j)}} \left[\frac{\partial \phi_{,j}}{\partial x_j} \right]_{,j}, \tag{22}$$

which gives an explicit form of the *fractal Laplacian*.

(iv) The curl of the curl operating on a vector field \mathbf{f} :

$$\text{curl} \times (\text{curl} \mathbf{f}) = \mathbf{e}_p e_{prj} \nabla_r^D (e_{jki} \nabla_k^D f_i) = \mathbf{e}_p \nabla_r^D \left(\nabla_p^D f_r \right) - \mathbf{e}_p \nabla_r^D \nabla_r^D f_p. \quad (23)$$

Next, the Helmholtz decomposition for fractals can be proved just as in the conventional case: a vector field \mathbf{F} with known divergence and curl, none of which equal to zero and which is finite and uniform and vanishes at infinity, may be expressed as the sum of a lamellar vector \mathbf{U} and a solenoidal vector \mathbf{V}

$$\mathbf{F} = \mathbf{U} + \mathbf{V} \quad (24)$$

with the operations

$$\text{curl} \mathbf{U} = \mathbf{0}, \quad \text{div} \mathbf{V} = 0 \quad (25)$$

understood in the sense of (18) and (19), respectively.

These results have recently been used (Ostojia-Starzewski 2012) to obtain Maxwell equations modified to generally anisotropic fractal media using two independent approaches: a conceptual one (involving generalized Faraday and Ampère laws) and the one directly based on a variational principle for electromagnetic fields. In both cases the resulting equations are the same, thereby providing a self-consistent verification of our derivations. Just as Tarasov (2010), we have found that the presence of anisotropy in the fractal structure leads to a source/disturbance as a result of generally unequal fractal dimensions in various directions, although, in the case of isotropy, our modified Maxwell equations are different. For most recent developments, see Tarasov (2014, 2015a,b).

Homogenization Process for Fractal Media

The formula (3) for fractal mass expresses the mass power law using fractional integrals. From a homogenization standpoint, this relationship allows an interpretation of the fractal medium as an intrinsically discontinuous continuum with a fractal metric embedded in the equivalent homogenized continuum model as shown in Fig. 2. In this figure, dl_{α_i} , dS_d , and dV_D represent the line, surface, and volume elements in the fractal medium, while dx_i , dS_2 , and dV_3 , respectively, denote these elements in the homogenized continuum model. The coefficients $c_1^{(i)}$, $c_2^{(k)}$, and c_3 provide the relationship between the fractal medium and the homogenized continuum model:

$$dl_{\alpha_i} = c_1^{(i)} dx_i, \quad dS_d = c_2^{(k)} dS_2, \quad dV_D = c_3 dV_3 \quad (\text{no sum}). \quad (26)$$

Standard image analysis techniques (such as the “box method” or the “sausage method” Stoyan and Stoyan 1994) allow a quantitative calibration of these coefficients for every direction and every cross-sectional plane. In a non-fractal medium

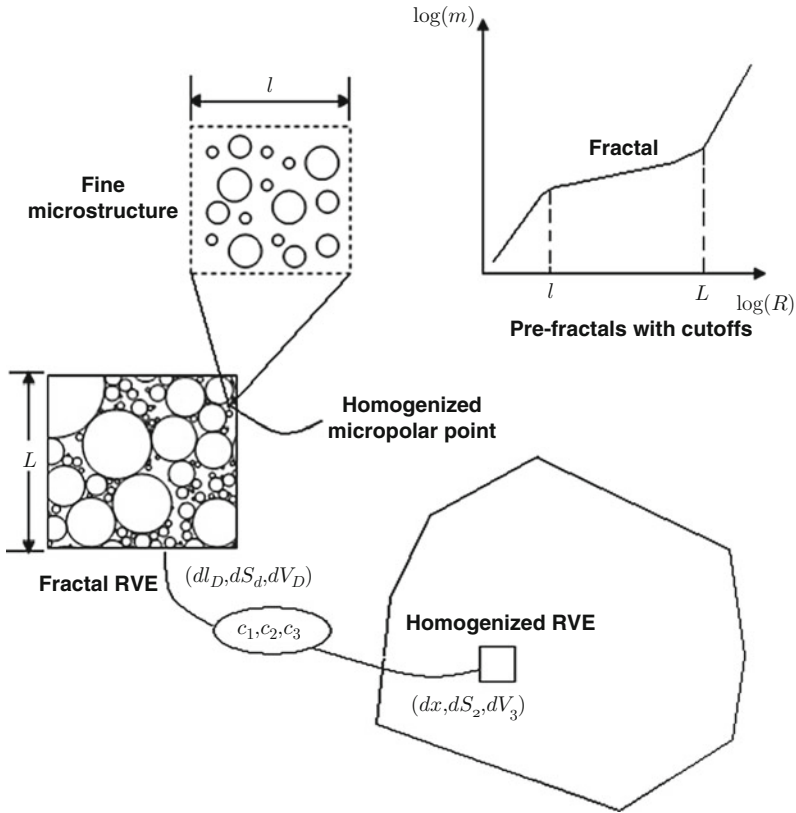


Fig. 2 Illustration of the two-level homogenization processes: fractal effects are present between the resolutions l and L in a fractal RVE with an Apollonian packing porous microstructure

where all the c coefficients in (26) are unity, one recovers conventional forms of the transport and balance equations of continuum mechanics. As discussed in section “[Fractal Angular Momentum Equation](#),” the presence of fractal geometric anisotropy ($c_1^{(j)} \neq c_1^{(k)}$, $j \neq k$, in general), as reflected by differences between the α ’s, leads to micropolar effects; see also Li and Ostoja-Starzewski (2010, 2011).

The above formulations provide one choice of calculus on fractals, i.e., through product integral (3), reflecting the mass scaling law (2) of fractal media. The advantage of our approach is that it is connected with conventional calculus through coefficients c_1, c_2 , and c_3 and therefore well suited for development of continuum mechanics and partial differential equations on fractal media as we shall see in the next sections. Besides, the product formulation allows a decoupling of coordinate variables, which profoundly simplifies the Gauss theorem (14) and many results thereafter. Other choices of calculus on fractals have been discussed in Ostoja-Starzewski et al. (2014).

Continuum Mechanics of Fractal Media

In the preceding section, we discussed product measures and fractional integrals, generalized the Gauss and Reynolds theorems to fractal media, and introduced fractal derivatives. We now have the framework to develop continuum mechanics in a fractal setting. We proceed just like it in classical continuum mechanics but employ fractional integrals expressed in terms of fractal derivatives.

In light of the discussion in section “[Homogenization Process for Fractal Media](#)” of the difference between fractal media and classical continuum mechanics, the definitions of stress and strain are to be modified appropriately. First, with reference to the modified concept of surface elements, we have the Cauchy tetrahedron of Fig. 3. Next, we specify the relationship between surface force, $\mathbf{F}^S (= F_k^S)$, and the Cauchy stress tensor $\boldsymbol{\sigma} (= \sigma_{kl})$ using fractional integrals as

$$F_k^S = \int_S \sigma_{lk} n_l dS_d, \tag{27}$$

where n_l are the components of the outward normal \mathbf{n} to S . On account of (26)₂, this force becomes

$$F_k^S = \int_S \sigma_{lk} n_l dS_d = \int_S \sigma_{lk} n_l c_2^{(l)} dS_2. \tag{28}$$

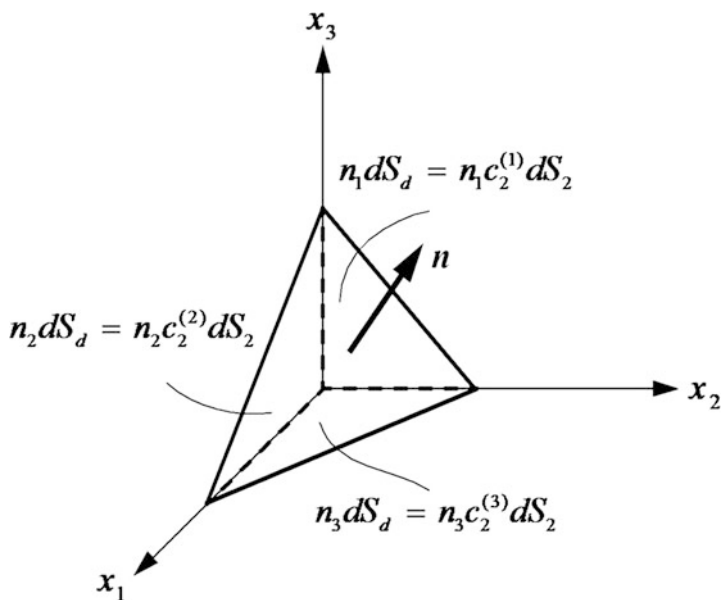


Fig. 3 Cauchy’s tetrahedron of a fractal body interpreted via product measures

To specify the strain, we observe, again using (26)₁ and the definition of fractal derivative (13), that

$$\frac{\partial}{\partial l_{\alpha_k}} = \frac{\partial x_k}{\partial l_{\alpha_k}} \frac{\partial}{\partial x_k} = \frac{1}{c_1^{(k)}} \frac{\partial}{\partial x_k} = \nabla_k^D. \quad (29)$$

Thus, for small deformations, we define the strain, ε_{ij} , in terms of the displacement u_k as

$$\varepsilon_{ij} = \frac{1}{2} \left(\nabla_j^D u_i + \nabla_i^D u_j \right) = \frac{1}{2} \left[\frac{1}{c_1^{(j)}} u_{i,j} + \frac{1}{c_1^{(i)}} u_{j,i} \right] \quad (\text{no sum}). \quad (30)$$

As shown in Li and Ostoja-Starzewski (2010), this definition of strain results in the same equations governing wave motion in linear elastic materials when derived by a variational approach as when derived by a mechanical approach, also see section “Fractal Wave Equations.”

In the following, we apply the balance laws for mass, linear momentum, energy, and entropy production to the fractal medium in order to derive the corresponding continuity equations.

Fractal Continuity Equation

Consider the equation for conservation of mass for \mathcal{W}

$$\frac{d}{dt} \int_{\mathcal{W}} \rho dV_D = 0, \quad (31)$$

where ρ is the density of the medium. Using the fractional Reynolds transport theorem (17), since \mathcal{W} is arbitrary, we find, in terms of the fractal derivative (13),

$$\frac{d\rho}{dt} + \rho c_1^{(k)} \nabla_k^D v_k = 0. \quad (32)$$

Fractal Linear Momentum Equation

Consider the balance law of linear momentum for \mathcal{W} ,

$$\frac{d}{dt} \int_{\mathcal{W}} \rho \mathbf{v} dV_D = \mathbf{F}^B + \mathbf{F}^S, \quad (33)$$

where \mathbf{F}^B is the body force, and \mathbf{F}^S is the surface force given by (28). In terms of the components of velocity, v_k , and body force density, X_k , (33) can be written as

$$\frac{d}{dt} \int_{\mathcal{W}} \rho v_k dV_D = \int_{\mathcal{W}} X_k dV_D + \int_{\partial\mathcal{W}} \sigma_{lk} n_l dS_d. \tag{34}$$

Using the Reynolds transport theorem and the continuity equation (32), the left-hand side is changed to

$$\begin{aligned} \frac{d}{dt} \int_{\mathcal{W}} \rho v_k dV_D &= \int_{\mathcal{W}} \left[\frac{\partial \rho v_k}{\partial t} + (v_k v_l \rho)_{,l} \right] dV_D \\ &= \int_{\mathcal{W}} \rho \left[\frac{\partial v_k}{\partial t} + v_l v_{k,l} \right] dV_D = \int_{\mathcal{W}} \rho \frac{dv_k}{dt} dV_D. \end{aligned} \tag{35}$$

Next, by the fractal Gauss theorem (14) and localization, we obtain the *fractal linear momentum equation*

$$\rho \dot{v}_k = X_k + \nabla_l^D \sigma_{lk}. \tag{36}$$

Fractal Angular Momentum Equation

The conservation of angular momentum in a fractal medium is stated as

$$\frac{d}{dt} \int_{\mathcal{W}} \rho e_{ijk} x_j v_k dV_D = \int_{\mathcal{W}} e_{ijk} x_j X_k dV_D + \int_{\partial\mathcal{W}} e_{ijk} x_j \sigma_{lk} n_l dS_d. \tag{37}$$

Using (37) and (14) yields

$$e_{ijk} \frac{\sigma_{jk}}{c_1^{(j)}} = 0. \tag{38}$$

It was shown in Li and Ostoja-Starzewski (2011) and Ostoja-Starzewski et al. (2016) that the presence of an anisotropic fractal structure is reflected by differences in the fractal dimensions α_i in different directions, which implies that $c_1^{(j)} \neq c_1^{(k)}$, $j \neq k$, in general. Therefore, the Cauchy stress is generally asymmetric in fractal media, indicating that the micropolar effects should be accounted for and (38) should be augmented by the presence of couple stresses. It is important to note here that a material may have anisotropic fractal structure yet be isotropic in terms of its constitutive laws (Joumaa and Ostoja-Starzewski 2011).

Focusing now on physical fractals (so-called pre-fractals), we consider a body that obeys a fractal mass power law (4) between the lower and upper cutoffs. The choice of the continuum approximation is specified by the resolution R . Choosing the upper cutoff, we arrive at the fractal representative volume element (RVE) involving a region up to the upper cutoff, which is mapped onto a homogenized continuum element in the whole body. The micropolar point homogenizes the very

fine microstructures into a rigid body (with 6 degrees of freedom) at the lower cutoff. The two-level homogenization processes are illustrated in Fig. 2.

To determine the inertia tensor \mathbf{i} at any micropolar point, we consider a rigid particle p having a volume \mathcal{P} , whose angular momentum is

$$\sigma_A = \int_{\mathcal{P}} (\mathbf{x} - \mathbf{x}_A) \times \mathbf{v}(\mathbf{x}, t) d\mu(\mathbf{x}). \quad (39)$$

Taking $\mathbf{v}(\mathbf{x}, t)$ as a helicoidal vector field (for some vector $\boldsymbol{\omega} \in \mathbb{R}^3$), $\mathbf{v}(\mathbf{x}, t) = \mathbf{v}(\mathbf{x}_A, t) + \boldsymbol{\omega} \times (\mathbf{x} - \mathbf{x}_A)$, we have found (Ostoja-Starzewski et al. 2016) all (diagonal and off-diagonal) components of i_{kl} as

$$I_{kl} = \int_{\mathcal{P}} [x_m x_m \delta_{kl} - x_k x_l] \rho(\mathbf{x}) dV_D. \quad (40)$$

Here we use I_{kl} (Eringen uses i_{kl}) in the current state so as to distinguish it from I_{KL} in the reference state, the relation between both being given by Eringen (1999) $I_{kl} = I_{KL} \chi_{kK} \chi_{lL}$, where χ_{kK} is a microdeformation tensor, or deformable director. For the entire fractal particle \mathcal{P} , we have

$$\frac{d}{dt} \int_{\mathcal{P}} \rho I_{KL} dV_D = 0, \quad (41)$$

which, in view of (17), results in the *fractal conservation of microinertia*

$$\frac{d}{dt} I_{kl} = I_{kr} v_{lr} + I_{lr} v_{kr}. \quad (42)$$

In micropolar continuum mechanics (Nowacki 1986; Eringen 1999), one needs a couple-stress tensor $\boldsymbol{\mu}$ and a rotation vector $\boldsymbol{\varphi}$ augmenting, respectively, the Cauchy stress tensor $\boldsymbol{\tau}$ (thus denoted so as to distinguish it from the symmetric $\boldsymbol{\sigma}$) and the deformation vector \mathbf{u} . The surface force and surface couple in the fractal setting can be specified by fractional integrals of $\boldsymbol{\tau}$ and $\boldsymbol{\mu}$, respectively, as

$$T_k^S = \int_{\partial\mathcal{W}} \tau_{ik} n_i dS_d, \quad M_k^S = \int_{\partial\mathcal{W}} \mu_{ik} n_i dS_d. \quad (43)$$

The above is consistent with the relation of force tractions and couple tractions to the force stresses and couple stresses on any surface element dS_d

$$t_k = \tau_{ik} n_i, \quad m_k = \mu_{ik} n_i. \quad (44)$$

Now, proceeding in a fashion similar as before, we obtain (36) and the *fractal angular momentum equation*

$$\frac{e_{ijk}}{c_1^{(j)}} \tau_{jk} + \nabla_j^D \mu_{ji} + Y_i = I_{ij} \dot{w}_j. \tag{45}$$

In the above, Y_i is the body force couple, while $v_k (= \dot{u}_k)$ and $w_k (= \dot{\phi}_k)$ are the deformation and rotation velocities, respectively.

Fractal Energy Equation

Globally, the conservation of energy has the following form

$$\frac{d}{dt} \int_{\mathcal{W}} (e + k) dV_D = \int_{\mathcal{W}} (X_i v_i + Y_i w_i) dV_D + \int_{\partial \mathcal{W}} (t_i v_i + m_i w_i - q_i n_i) dS_d, \tag{46}$$

where $k = \frac{1}{2} (\rho v_i v_i + I_{ij} w_i w_j)$ is the kinetic energy density, e the internal energy density, and $\mathbf{q} (= q_i)$ heat flux through the boundary of \mathcal{W} . As an aside we note that, just like in conventional (non-fractal media) continuum mechanics, the balance equations of linear momentum (36) and angular momentum (47) can be consistently derived from the invariance of energy (48) with respect to rigid body translations ($v_i \rightarrow v_i + b_i$, $w_i \rightarrow w_i$) and rotations ($v_i \rightarrow v_i + e_{ijk} x_j \omega_k$, $w_i \rightarrow w_i + \omega_i$), respectively.

To obtain the expression for the rate of change of internal energy, we start from

$$\int_{\mathcal{W}} (\dot{e} + \rho v_i \dot{v}_i + I_{ij} w_i \dot{w}_j) dV_D = \int_{\mathcal{W}} [X_i v_i + Y_i w_i + \nabla_j^D (\tau_{ji} v_j + \mu_{ji} w_j)] dV_D - \int_{\mathcal{W}} \nabla_i^D q_i dV_D, \tag{47}$$

and note (36) and (47), to find

$$\dot{e} = \tau_{ji} \left(\nabla_j^D v_i - e_{kji} \omega_k \right) + \mu_{ji} \nabla_j^D w_i - \nabla_i^D q_i. \tag{48}$$

Next, introducing the infinitesimal strain tensor and the curvature tensor in fractal media

$$\gamma_{ji} = \nabla_j^D u_i - e_{kji} \varphi_k, \quad \kappa_{ji} = \nabla_j^D \varphi_i, \tag{49}$$

we find that the energy balance (52) can be written as

$$\dot{e} = \tau_{ij} \dot{\gamma}_{ij} + \mu_{ij} \dot{\kappa}_{ij}. \tag{50}$$

Assuming e to be a state function of γ_{ij} and κ_{ij} only and assuming τ_{ij} and μ_{ij} not to be explicitly dependent on the temporal derivatives of γ_{ij} and κ_{ij} , we find

$$\tau_{ij} = \frac{\partial e}{\partial \gamma_{ij}} \quad \mu_{ij} = \frac{\partial e}{\partial \kappa_{ij}}. \quad (51)$$

which shows that, just like in non-fractal continuum mechanics, also in the fractal setting, (τ_{ij}, γ_{ij}) and (μ_{ij}, κ_{ij}) are conjugate pairs.

Fractal Second Law of Thermodynamics

To derive the field equation of the second law of thermodynamics in a fractal medium $B(\omega)$, we begin with the global form of that law in the volume V_D , having a Euclidean boundary $\partial \mathcal{W}$, that is

$$\dot{S} = \dot{S}^{(r)} + \dot{S}^{(i)} \quad \text{with} \quad \dot{S}^{(r)} = \frac{\dot{Q}}{T}, \quad \dot{S}^{(i)} \geq 0, \quad (52)$$

where \dot{S} , $\dot{S}^{(r)}$ and $\dot{S}^{(i)}$ stand, respectively, for the total, reversible, and irreversible entropy production rates in V_D . Equivalently,

$$\dot{S} \geq \dot{S}^{(r)}. \quad (53)$$

Since these two rates are extensive quantities, we obtain

$$\int_W \rho \frac{d}{dt} s \, dV_D = \frac{d}{dt} \int_W \rho s \, dV_D \geq - \int_{\partial W} \frac{q_k n_k}{T} \, dS_d = - \int_W \nabla_k^D \left(\frac{q_k}{T} \right) \, dV_D, \quad (54)$$

which yields the local form of the second law

$$\rho \frac{ds}{dt} \geq - \nabla_k^D \left(\frac{q_k}{T} \right) - \frac{\nabla_k^D q_k \cdot k}{T} + \frac{q_k \nabla_k^D T}{T^2}. \quad (55)$$

Just like in thermomechanics of non-fractal bodies (Ziegler 1983; Maugin 1999), we now introduce the rate of irreversible entropy production $\rho \dot{s}^{(i)}$ which, in view of (57), gives

$$0 \leq \rho \dot{s}^{(i)} = \rho \dot{s} + \frac{\nabla_k^D q_k \cdot k}{T} - \frac{q_k \nabla_k^D T}{T^2} + \rho h. \quad (56)$$

Here with s we denote specific entropies (i.e., per unit mass). Next, we recall the classical relation between the free energy density ψ , the internal energy density e , the entropy s , and the absolute temperature T : $\psi = e - Ts$. This allows us to write for time rates of these quantities $\dot{\psi} = \dot{e} - s\dot{T} - T\dot{s}$. On the other hand, with ψ being a function of the strain γ_{ji} and curvature-torsion κ_{ji} tensors, the internal variables α_{ij} (strain type) and ζ_{ij} (curvature-torsion type), and temperature T , we have

$$\rho \dot{\psi} = \rho \frac{\partial \psi}{\partial \gamma_{ij}} \dot{\gamma}_{ij} + \rho \frac{\partial \psi}{\partial \alpha_{ij}} \dot{\alpha}_{ij} + \rho \frac{\partial \psi}{\partial \kappa_{ij}} \dot{\kappa}_{ij} + \rho \frac{\partial \psi}{\partial \zeta_{ij}} \dot{\zeta}_{ij} + \rho \frac{\partial \psi}{\partial T} \dot{T}. \tag{57}$$

In the above we have adopted the conventional relations giving the (external and internal) quasi-conservative Cauchy and Cosserat (couple) stresses as well as the entropy density as gradients of ψ

$$\tau_{ij}^{(q)} = \rho \frac{\partial \psi}{\partial \varepsilon_{ij}} \beta_{ij}^{(q)} = \rho \frac{\partial \psi}{\partial \alpha_{ij}} \mu_{ij}^{(q)} = \rho \frac{\partial \psi}{\partial \kappa_{ij}} \eta_{ij}^{(q)} = \rho \frac{\partial \psi}{\partial \zeta_{ij}} s = - \frac{\partial \psi}{\partial T} \tag{58}$$

This is accompanied by a split of total Cauchy and micropolar stresses into their quasi-conservative and dissipative parts

$$\tau_{ij} = \tau_{ij}^{(q)} + \tau_{ij}^{(d)}, \quad \mu_{ij} = \mu_{ij}^{(q)} + \mu_{ij}^{(d)}, \tag{59}$$

along with relations between the internal quasi-conservative and dissipative stresses

$$\beta_{ij}^{(q)} = -\beta_{ij}^{(d)}, \quad \eta_{ij}^{(q)} = -\eta_{ij}^{(d)}. \tag{60}$$

In view of (59) and (60), we obtain

$$\rho (\dot{\psi} + s\dot{T}) = \rho (\dot{e} - T\dot{s}) = \tau_{ij}^{(q)} \dot{\gamma}_{ij} + \beta_{ij}^{(q)} \dot{\alpha}_{ij} + \mu_{ij}^{(q)} \dot{\kappa}_{ij} + \eta_{ij}^{(q)} \dot{\zeta}_{ij} + \rho h. \tag{61}$$

On account of the energy balance, this is equivalent to

$$T \rho \dot{s} = \tau_{ij}^{(d)} \dot{\gamma}_{ij} + \beta_{ij}^{(d)} \dot{\alpha}_{ij} + \mu_{ij}^{(d)} \dot{\kappa}_{ij} + \eta_{ij}^{(d)} \dot{\zeta}_{ij} - \nabla_k^D q_k + \rho h. \tag{62}$$

Recalling (61), we find the local form of the second law in terms of time rates of strains and internal parameters

$$0 \leq T \rho \dot{s}^{(i)} = \tau_{ij}^{(d)} \dot{\gamma}_{ij} + \beta_{ij}^{(d)} \dot{\alpha}_{ij} + \mu_{ij}^{(d)} \dot{\kappa}_{ij} + \eta_{ij}^{(d)} \dot{\zeta}_{ij} - \frac{q_k \nabla_k^D T}{T} + \rho h. \tag{63}$$

The above is a generalization of the Clausius-Duhem inequality to fractal dissipative media with internal parameters. Upon dropping the internal parameters (which may well be the case for a number of materials), the terms $\beta_{ij}^{(d)} \dot{\alpha}_{ij}$ and $\eta_{ij}^{(d)} \dot{\zeta}_{ij}$ drop out, whereas, upon neglecting the micropolar effects, the terms $\mu_{ij}^{(d)} \dot{\kappa}_{ij}$ and $\eta_{ij}^{(d)} \dot{\zeta}_{ij}$ drop out.

For non-fractal bodies without internal parameters, the stress tensor τ_{ij} reverts back to σ_{ij} , and (63) reduces to the simple well-known form Ziegler (1983)

$$0 \leq T \rho \dot{s}^{(i)} = \sigma_{ij}^{(d)} \dot{\gamma}_{ij} - \frac{T_{,k} q_k}{T} + \rho h. \tag{64}$$

Interestingly, the fractal gradient (13) appears only in the thermal dissipation term. In fact, this derivative arises in processes of heat transfer in a fractal rigid conductor and coupled thermoelasticity of fractal deformable media (Oldham and Spanier 1974).

Fractal Wave Equations

Just like in conventional continuum mechanics, the basic continuum equations for fractal media presented above have to be augmented by constitutive relations. At this point, we expect that the fractal geometry influences configurations of physical quantities like stress and strain, but does not affect the physical laws (like conservation principles) and constitutive relations that are inherently due to material properties. This expectation is supported in two ways:

- A study of scale effects of material strength and stress from the standpoint of fractal geometry which is confirmed by experiments involving both brittle and plastic materials (Carpinteri and Pugno 2005)
- Derivations of wave equations through mechanical and variational approaches, respectively (Demmie and Ostoja-Starzewski 2011), as shown below

Elastodynamics of a Fractal Timoshenko Beam

Analogous results, also exhibiting self-consistency, were obtained in elastodynamics of a fractally structured Timoshenko beam (Li and Ostoja-Starzewski 2009a). First recall that such a beam model has two degrees of freedom (q_1, q_2) at each point: the transverse displacement $q_1 = w$ and the rotation $q_2 = \varphi$. In the mechanical approach, the beam equation can be derived from the force and moment balance analysis. Thus, beginning with the expressions of shear force (V) and bending moment (M)

$$V = \kappa\mu A (\nabla_x^D w - \varphi), \quad M = -EI \nabla_x^D \varphi, \quad (65)$$

we find

$$\rho A \ddot{w} = \nabla_x^D V, \quad \rho I \ddot{\varphi} = V - \nabla_x^D M. \quad (66)$$

which lead to

$$\begin{aligned} \rho A \ddot{w} &= \nabla_x^D [\kappa\mu A (\nabla_x^D w - \varphi)], \\ \rho I \ddot{\varphi} &= \nabla_x^D (EI \nabla_x^D \varphi) + \kappa\mu A (\nabla_x^D w - \varphi). \end{aligned} \quad (67)$$

The kinetic energy is

$$T = \frac{1}{2} \rho \int_0^l \left[I (\dot{\varphi})^2 + A (\dot{w})^2 \right] dl_D, \tag{68}$$

while the potential energy is

$$\begin{aligned} U &= \frac{1}{2} \int_0^l \left[EI \left(\frac{\partial \varphi}{\partial l_D} \right)^2 + \kappa \mu A \left(\frac{\partial w}{\partial l_D} - \varphi \right)^2 \right] dl_D \\ &= \frac{1}{2} \int_0^l \left[EI c_1^{-2} (\varphi_{,x})^2 + \kappa \mu A (c_1^{-1} w_{,x} - \varphi)^2 \right] c_1 dx. \end{aligned} \tag{69}$$

Now, the Euler-Lagrange equations

$$\frac{\partial}{\partial t} \left[\frac{\partial L}{\partial \dot{q}_i} \right] + \sum_{j=1}^3 \frac{\partial}{\partial x_j} \left[\frac{\partial L}{\partial (q_{i,j})} \right] - \frac{\partial L}{\partial q_i} = 0 \tag{70}$$

result in the same as above.

In the case of elastostatics and when the rotational degree of freedom ceases to be independent ($\varphi = \partial w / \partial l_D = \nabla_x^D w$), we find the equation of a fractal Euler-Bernoulli beam

$$\nabla_x^D \nabla_x^D (EI \nabla_x^D \nabla_x^D w) = 0, \tag{71}$$

which shows that

$$M = EI \nabla_x^D \nabla_x^D w. \tag{72}$$

The relationship between the bending moment (M) and the curvature ($\nabla_x^D \nabla_x^D w$) still holds, while c_1 enters the determination of curvature ($\nabla_x^D \nabla_x^D w = c_1^{-1} (c_1^{-1} w_{,x})_{,x}$).

In a nutshell, the fractional power law of mass implies a fractal dimension of scale measure, so the derivatives involving spatial scales should be modified to incorporate such effect by postulating c_1, c_2 , and c_3 coefficients, according to the material body being embedded in a 1D, 2D, or 3D Euclidean space.

More work was done on waves in linear elastic fractal solids under small motions. Several cases of isotropic (Joumaa and Ostoja-Starzewski 2011) or anisotropic (with micropolar effects) (Joumaa et al. 2014; Joumaa and Ostoja-Starzewski 2016) media have been considered through analytical and computational methods. It was found on the mathematical side that fractal versions of harmonic, Bessel, and Hankel functions. The study (Joumaa and Ostoja-Starzewski 2011) provided the basis for a study of wave motion in a human head, where the brain, while

protected by skull bones, is actually surrounded by a cerebrospinal fluid (so-called CSF) and has fractal geometric characteristics (Joumaa and Ostoja-Starzewski 2013).

Elastodynamics in 3D

First, the equations of motion of 3D elastodynamics can be determined for finite strain motions by extending the procedure of classical continuum mechanics. To this end, begin with the action functional of a fractal solid \mathcal{W} isolated from external interactions

$$\delta\mathcal{I} = \delta \int_{t_1}^{t_2} [\mathcal{K} - \mathcal{E}] dt = 0, \quad (73)$$

where \mathcal{K} and \mathcal{E} are the kinetic and internal energies

$$\mathcal{K} = \frac{1}{2} \int_{\mathcal{W}} \rho v_i v_i dV_D \quad \mathcal{E} = \int_{\mathcal{W}} \rho e dV_D. \quad (74)$$

Thus, we have a functional, which can be rewritten in fractal space-time (e being the specific, per unit mass, internal energy density) as

$$0 = \delta\mathcal{I} = \delta \int_{t_1}^{t_2} \int_{\mathcal{W}} \left[\frac{1}{2} \rho v_k^2 - \rho e \right] dV_D dt = \delta \int_{t_1}^{t_2} \int_{\mathcal{W}} \left[\rho \frac{1}{2} v_k^2 - \rho e \right] dV_D dt. \quad (75)$$

Analogous to the strain of (30), the deformation gradient is

$$F_{kI} = \nabla_I^D x_k \equiv \frac{1}{c_1^{(I)}} x_{k,I}. \quad (76)$$

Assuming that the specific energy density e depends only on the deformation gradient, and that ρ has no explicit dependence on time, leads to:

(i) The boundary conditions

$$\partial\mathcal{W} \rho \frac{\partial e}{\partial F_{kI}} N_I = 0 \quad \text{or} \quad \delta x_k = 0 \quad \text{on} \quad \partial\mathcal{W} \quad (77)$$

(ii) The kinematic constraints $\delta x_k = 0$ at $t = t_1$ and $t = t_2$

The Hamilton principle for (73) implies the equation governing motion in a fractal solid under finite strains

$$\nabla_I^D \left[\rho \frac{\partial e}{\partial F_{kI}} \right] - \rho \frac{dv_k}{dt} = 0. \quad (78)$$

Next, in fractal bodies without internal dissipation, e plays the role of a potential

$$T_{kI} = \rho \frac{\partial e}{\partial F_{kI}}, \tag{79}$$

where T_{kI} is the first Piola-Kirchhoff stress tensor, and (78) becomes

$$\nabla_I^D T_{kI} - \rho \frac{dv_k}{dt} = 0 \quad \text{or} \quad \frac{1}{c_1^{(I)}} T_{kI,I} - \rho \frac{dv_k}{dt} = 0. \tag{80}$$

Restricting the motion to small deformation gradients, T_{kI} becomes the Cauchy stress tensor, and we recover the linear momentum equation (37). Generalizing this to the situation of \mathcal{W} interacting with the environment (i.e., subject to body forces), $\delta(\mathcal{I} + \mathcal{W} - \mathcal{P}) = 0$, we obtain

$$\nabla_I^D T_{kI} + \rho \left(b_k - \frac{dv_k}{dt} \right) = 0. \tag{81}$$

The above formulation was specialized in Demmie and Ostoja-Starzewski (2011) to 1D models and focused on nonlinear waves, demonstrating how the governing equations might be solved by the method of characteristics in fractal space-time. We also studied shock fronts in linear viscoelastic solids under small strains. We showed that the discontinuity in stress across a shock front in a fractal medium is identical to the classical result.

Related Topics

Extremum and Variational Principles in Fractal Bodies

Just like in preceding sections, the dimensional regularization approach can also be applied to other statements in continuum/solid mechanics involving integral relations. For example, the Maxwell-Betti reciprocity relation of linear elasticity $\int_{\partial W} t_i^* u_i dS_2 = \int_{\partial W} t_i u_i^* dS_2$ is generalized for fractal media to

$$\int_{\partial W} t_i^* u_i dS_d + \int_{\partial W} m_i^* \varphi_i dS_d = \int_{\partial W} t_i u_i^* dS_d + \int_{\partial W} m_i \varphi_i^* dS_d, \tag{82}$$

so as to read in \mathbb{E}^3 :

$$\int_{\partial W} t_i^* u_i c_2 dS_2 + \int_{\partial W} m_i^* \varphi_i c_2 dS_2 = \int_{\partial W} t_i u_i^* c_2 dS_2 + \int_{\partial W} m_i \varphi_i c_2 dS_2. \tag{83}$$

The reciprocity relation (82) is proved by appealing to the Green-Gauss theorem and the Hooke law ($\sigma_{ij} = C_{ijkl} \varepsilon_{kl}$) and proceeding just like in the conventional

continuum elasticity. As an application, consider the classical problem of calculation of the reduction in volume ΔV of a linear elastic isotropic body (of bulk modulus κ) due to two equal, collinear, opposite forces F , separated by a distance L . Clearly, in the classical case, one does not need the micropolar term and, as discussed in Li and Ostoja-Starzewski (2009a), finds $\Delta V = FL/3\kappa$. On the other hand, for a fractal body, given the loading $\tau_{ij} = -p\delta_{ij}$ and $\mu_{ij} = 0$, the integrals involving scalar products of couple traction with rotation vanish, and we obtain

$$p \Delta V c_3 = p \frac{FL}{3\kappa} c_1. \quad (84)$$

Recalling (6) and assuming that the opposite forces are applied parallel to one of the axis of the coordinate system (say, x_k) yield $\Delta V = FL/3\kappa c_2^{(k)}$. This can be evaluated numerically for a specific material according to (6).

Next, suppose we focus on situations where couple-stress effects are negligible. Then, we recall the concept of a *statically admissible field* as a tensor function $\sigma_{ij}(\mathbf{x})$, such that $\sigma_{ij} = \sigma_{ji}$ and $(F_k = \rho f_k)$

$$F_k + \nabla_l^D \sigma_{kl} = 0 \quad (85)$$

in W and the boundary conditions

$$\sigma_{kl} n_l = t_k \quad (86)$$

on ∂W_t . Similarly, we recall a *kinematically admissible displacement field* as a vector function $\mathbf{u}(\mathbf{x})$ satisfying the boundary conditions

$$u_i = f_i \quad (87)$$

on ∂W_u . We can now consider the *principle of virtual work*: “The virtual work of the internal forces equals the virtual work of the external forces.” Let $\boldsymbol{\sigma}(\mathbf{x})$ be a statically admissible stress field and $\mathbf{u}(\mathbf{x})$ a kinematically admissible displacement field. Define $\varepsilon_{ij}(\mathbf{u}) = u_{(i,j)}$. Then

$$\int_W \sigma_{ij} \varepsilon_{ij} dV_D = \int_W F_i u_i dV_D + \int_{\partial W_t} t_i u_i dS_d + \int_{\partial W_u} \sigma_{ij} n_j f_i dS_d. \quad (88)$$

The proof follows by substitution from the fractional equation of static equilibrium and boundary conditions after integrating by parts, and using the Gauss theorem, all conducted over the fractal domain W . Of course, the above can be rewritten in terms of conventional integrals:

$$\begin{aligned} \int_W \sigma_{ij} \varepsilon_{ij} c_3 dV_3 &= \int_W F_i u_i c_3 dV_3 + \int_{\partial W_t} c_2 t_i u_i c_2 dS_2 \\ &+ \int_{\partial W_t} \sigma_{ij} n_j f_i c_2 dS_2. \end{aligned} \quad (89)$$

In a similar way, we can adapt to fractal elastic bodies the principle of virtual displacement, principle of virtual stresses, principle of minimum potential energy, principle of minimum complementary energy, and related principles for elastic-plastic or rigid bodies.

Relation to other studies of complex systems: While the calculus in non-integer spaces outlined in this brief review has been employed for fractal porous media, the approach can potentially be extended to microscopically heterogeneous physical systems (made up of many different micromechanical systems, having different physical properties, and interacting under the influence of different kinds of phenomena). Several such possibilities have been discussed in Ostoja-Starzewski et al. (2014), and one paradigm is given in the next section.

Fracture in Elastic-Brittle Fractal Solids

General Considerations

According to Griffith’s theory of elastic-brittle solids, the strain energy release rate G is given by Gdoutos (1993)

$$G = \frac{\partial W}{\partial A} - \frac{\partial \mathcal{E}^e}{\partial A} = 2\gamma, \tag{90}$$

where A is the crack surface area formed, W is the work performed by the applied loads, \mathcal{E}^e is the elastic strain energy, and γ is the energy required to form a unit of new material surface. The material parameter γ is conventionally taken as constant, but, given the presence of a randomly microheterogeneous material structure, its random field nature is sometimes considered explicitly (Ostoja-Starzewski 2008b, 2004). Recognizing that the random material structure also affects the elastic moduli (such as E), the computation of \mathcal{E}^e and G in (90) needs to be reexamined (Ostoja-Starzewski 2004); see also Balankin et al. (2011) in the context of paper mechanics. With reference to Fig. 4, we consider a 3D material body described by D and d and having a crack of depth a and a fractal dimension DF .

Focusing on a fractal porous material, we have

$$U^e = \int_W \rho u dV_D = \int_W \rho u c_3 dV_3. \tag{91}$$

By revising Griffith’s derivation for a fractal elastic material, we then obtain

$$U^e = \frac{\pi a^2 c_1^2 \sigma^2}{8\mu} (K + 1) c_3, \tag{92}$$

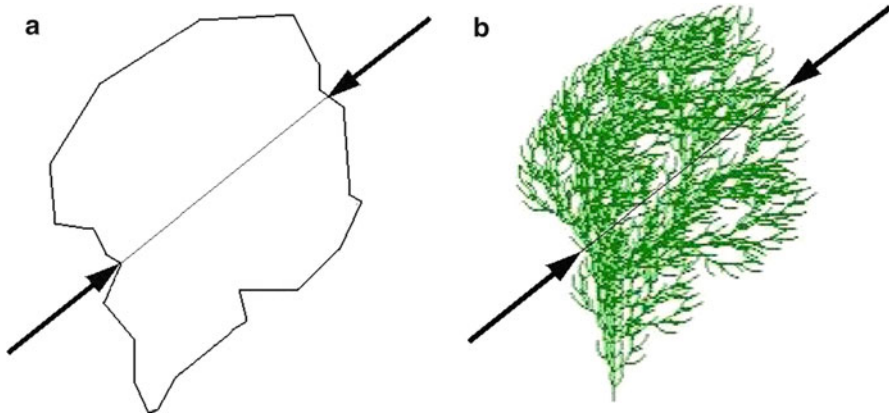


Fig. 4 (a) A fractal body subjected to two equal, collinear, opposite forces F and (b) its homogenized equivalent via dimensional regularization

with ν being the Poisson ratio, and

$$K = \begin{cases} 3 - 4\nu & \text{for plane strain} \\ \frac{3 - \nu}{1 + \nu} & \text{for plane stress} \end{cases} \quad (93)$$

the Kolosov constant.

Dead-load conditions. Equation (90) becomes

$$G = \frac{\partial U^e}{\partial A} = 2\gamma. \quad (94)$$

If $A = 2a \times 1$, this gives the critical stress

$$\sigma_c = \sqrt{\frac{2\gamma E}{(1 - \nu^2)\pi a c_3 c_1^2}}. \quad (95)$$

However, if the fracture surface is fractal of a dimension α , then we should use $\partial/\partial l_{DF}$ instead of $\partial/\partial a$. Now, since we have (note Fig. 5) $dl_{DF} = c_1 da$, the new partial derivative becomes

$$\frac{\partial}{\partial l_{DF}} = \frac{\partial}{c_1 \partial a}, \quad (96)$$

so that

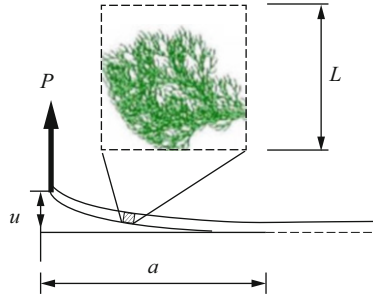


Fig. 5 Fracture and peeling of a microbeam of thickness L off a substrate. A representative volume element dV_3 imposed by the pre-fractal structure characterized by upper cutoff scale L is shown. Thus, the beam is homogeneous above the length scale L . By introducing random variability in that structure, one obtains a random beam; see (103)

$$\sigma_c = \sqrt{\frac{2\gamma E}{(1 - \nu^2)\pi a c_3 c_1}}. \tag{97}$$

Fixed-grip conditions. We consider the case of a crack of depth a and width B in plane strain. In this case the displacement is constant (i.e., non-random), and the load is random. Now, only the second term in (90) remains, so that

$$G = -\frac{\partial \mathcal{E}^e(a)}{B \partial l_{DF}} = -\frac{\partial \mathcal{E}^e(a)}{B c_1 \partial a}. \tag{98}$$

Peeling a Layer Off a Substrate

Dead-load conditions. As a specific case, we take an Euler-Bernoulli beam, so that the strain energy is

$$\mathcal{E}(a) = \int_0^a \frac{M^2}{2IE} dx, \tag{99}$$

where a is crack length, M is bending moment, and I is beam’s moment of inertia. Henceforth, we simply work with $a = A/B$, where B is the constant beam (and crack) width. In view of Clapeyron’s theorem, the strain energy release rate may be written as

$$G = \frac{\partial \mathcal{E}}{B \partial a}. \tag{100}$$

For a layer modeled as a fractal Euler-Bernoulli beam (section “[Peeling a Layer Off a Substrate](#)”), we have

$$\mathcal{E}(a) = \int_0^a \frac{M^2}{2IE} dl_D = \int_0^a \frac{M^2}{2IE} c_1 dx, \quad (101)$$

so that

$$G = \frac{\partial \mathcal{E}}{c_1 B \partial a}. \quad (102)$$

Generalization to a Statistical Ensemble

Now, if the beam's material is random, E is a random field parametrized by x , which we can write as a sum of a constant mean $\langle E \rangle$ and a zero-mean fluctuation $E'(x)$

$$E(\omega, x) = \langle E \rangle + E'(\omega, x) \quad \omega \in \Omega, \quad (103)$$

where ω is an elementary event in Ω , a sample space. Clearly, \mathcal{E} is a random integral, such that, for each and every realization $\omega \in \Omega$, we should consider

$$\mathcal{E}(a, E(\omega)) = \int_0^a \frac{M^2 c_1 dx}{2IE(\omega, x)}. \quad (104)$$

Upon ensemble averaging, this leads to an average energy

$$\langle \mathcal{E}(a, E) \rangle = \left\langle \int_0^a \frac{M^2 c_1 dx}{2I [\langle E \rangle + E'(\omega, x)]} \right\rangle. \quad (105)$$

In the conventional formulation of deterministic fracture mechanics, random microscale heterogeneities $E'(x, \omega)$ are disregarded, and (104) is evaluated by simply replacing the denominator by $\langle E \rangle$, so that

$$\mathcal{E}(a, \langle E \rangle) = \int_0^a \frac{M^2 c_1 dx}{2I \langle E \rangle}. \quad (106)$$

Clearly, this amounts to postulating that the response of an idealized homogeneous material is equal to that of a random one on average. To make a statement about $\langle \mathcal{E}(a, E) \rangle$ versus $\mathcal{E}(a, \langle E \rangle)$, and about $\langle G(E) \rangle$ versus $G(\langle E \rangle)$, first, note the random field E is positive valued almost surely. Then, Jensen's inequality yields a relation between harmonic and arithmetic averages of the random variable $E(\omega)$

$$\frac{1}{\langle E \rangle} \leq \left\langle \frac{1}{E} \right\rangle. \quad (107)$$

whereby the x -dependence is immaterial in view of the assumed wide-sense stationary of field E . With (105) and (106), and assuming that the conditions required by Fubini's theorem are met, this implies that

$$\begin{aligned} \mathcal{E}(a, \langle E \rangle) &= \int_0^a \frac{M^2 c_1 dx}{2I \langle E \rangle} \leq \int_0^a \frac{M^2 c_1}{2I} \left\langle \frac{1}{E} \right\rangle dx \\ &= \left\langle \int_0^a \frac{M^2 c_1 dx}{2IE(\omega, x)} \right\rangle = \langle \mathcal{E}(a, E) \rangle, \end{aligned} \tag{108}$$

Now, defining the strain energy release rate $G(a, \langle E \rangle)$ in a reference material specified by $\langle E \rangle$, and the strain energy release rate $\langle G(a, E) \rangle$ properly ensemble averaged in the random material $\{E(\omega, x); \omega \in \Omega, x \in [0, a]\}$

$$G(a, \langle E \rangle) = \frac{\partial \mathcal{E}(a, \langle E \rangle)}{B c_1 \partial a} \quad \langle G(a, E) \rangle = \frac{\partial \langle \mathcal{E}(a, E) \rangle}{B c_1 \partial a}, \tag{109}$$

and noting that the side condition is the same in both cases

$$\mathcal{E}(a, \langle E \rangle) |_{a=0} = 0 \quad \langle \mathcal{E}(a, E) \rangle |_{a=0} = 0, \tag{110}$$

we find

$$G(a, \langle E \rangle) \leq \langle G(a, E) \rangle. \tag{111}$$

This provides a formula for the ensemble average G under dead-load conditions using deterministic fracture mechanics for Euler-Bernoulli beams made of fractal random materials.

Just like in the case of non-fractal materials (Ostoj-Starzewski 2004), the inequality (111) shows that G computed under the assumption that the random material is directly replaced by a homogeneous material ($E(x, \omega) = \langle E \rangle$) is lower than G computed with E taken explicitly as a spatially varying material property.

Fixed-grip conditions. On account of (4), assuming that there is loading by a force P at the tip, we obtain

$$G = -\frac{u}{2B c_1} \frac{\partial P}{\partial a}. \tag{112}$$

Take now a cantilever beam problem implying $P = 3uEI/(c_1 a)^3$. Then, we find

$$\langle G \rangle = -\frac{u}{2B c_1} \left\langle \frac{\partial P}{\partial a} \right\rangle = -\frac{u}{2B c_1} \frac{\partial \langle P \rangle}{\partial a} = \frac{9u^2 I \langle E \rangle}{2B (c_1 a)^4}. \tag{113}$$

Since the load – be it a force and/or a moment – is always proportional to E , this indicates that G can be computed by direct ensemble averaging of E under fixed-grip loading, and, indeed, the same conclusion carries over to Timoshenko beams. This analysis may be extended to mixed-loading conditions and stochastic crack

stability by generalizing the study of non-fractal, random beams carried out in Li and Ostoja-Starzewski (2009a).

Closure

While fractals are abundant in nature, the mechanics of fractal media is still in its infancy. In this brief review, we showed how fundamental balance laws for fractal porous media can be developed using a homogenization method using calculus in non-integer dimensional spaces, a special version of dimensional regularization. The developments are confined to materials whose mass is specified by a product measure, where the overall mass fractal dimension D equals $\alpha_1 + \alpha_2 + \alpha_3$ along the diagonals, $|x_1| = |x_2| = |x_3|$, where each α_k plays the role of a fractal dimension in the direction x_k . The strategy is to express the balance laws for fractal media in terms of fractional integrals and, then, to convert them to integer-order integrals in conventional (Euclidean) space. This leads to balance laws expressed in continuous forms albeit modified by the presence of coefficients responsible for dependencies on α_1 's; this makes mathematical treatments more tractable. It is shown how to develop balance laws of fractal media (continuity, linear and angular momenta, first and second law of thermodynamics) as well as the elastodynamics equations of a fractal Timoshenko beam and 3D solids under finite strains. In general, if the fractal geometry is described by a product measure, the angular momentum balance cannot be satisfied unless the stress tensor is asymmetric. Therefore, a generally anisotropic fractal medium necessitates adoption of a nonclassical continuum, the Cosserat model, for the balance laws to be satisfied (Maugin 2016). We then discuss extremum and variational principles and fracture mechanics in fractal media. In all the cases, the derived equations for fractal media depend explicitly on fractal dimensions and reduce to conventional forms for continuous media with Euclidean geometries upon setting the dimensions to integers. The fractal model discussed is useful for solving complex mechanics problems involving fractal materials composed of microstructures of inherent length scale, generalizing the universally applied classical theory of elastodynamics for continuous bodies. This modeling approach offers a foundation upon which one can study (thermo)mechanical phenomena involving fractals analytically and computationally while determining fractal dimensions by image analyses.

Acknowledgements This work was made possible by the support from NSF (grant CMMI-1462749).

References

- A.S. Balankin, O. Susarrey, C.A. Mora Santos, J. Patino, A. Yagues, E.I. Garcia, Stress concentration and size effect in fracture of notched heterogeneous material. *Phys. Rev. E* **83**, 015101(R) (2011)

- M.F. Barnsley, *Fractals Everywhere* (Morgan Kaufmann, San Francisco, 1993)
- A. Carpinteri, B. Chiaia, P.A. Cornetti, A disordered microstructure material model based on fractal geometry and fractional calculus. *ZAMP* **84**, 128–135 (2004)
- A. Carpinteri, N. Pugno, Are scaling laws on strength of solids related to mechanics or to geometry? *Nat. Mater.* **4**, 421–23 (2005)
- P.N. Demmie, Ostoja-Starzewski, Waves in fractal media. *J. Elast.* **104**, 187–204 (2011)
- A.C. Eringen, *Microcontinuum Field Theories I* (Springer, New York, 1999)
- K. Falconer, *Fractal Geometry: Mathematical Foundations and Applications* (Wiley, Chichester, 2003).
- E.E. Gdoutos, *Fracture Mechanics: An Introduction* (Kluwer Academic Publishers, Dordrecht, 1993)
- H.M. Hastings, G. Sugihara, *Fractals: A User's Guide for the Natural Sciences* (Oxford Science Publications, Oxford, 1993)
- H. Joumaa, M. Ostoja-Starzewski, On the wave propagation in isotropic fractal media. *ZAMP* **62**, 1117–1129 (2011)
- H. Joumaa, M. Ostoja-Starzewski, Acoustic-elastodynamic interaction in isotropic fractal media. *Eur. Phys. J. Spec. Top.* **222**, 1949–1958 (2013)
- H. Joumaa, M. Ostoja-Starzewski, P.N. Demmie, Elastodynamics in micropolar fractal solids. *Math. Mech. Solids* **19**(2), 117–134 (2014)
- H. Joumaa, M. Ostoja-Starzewski, On the dilatational wave motion in anisotropic fractal solids. *Math. Comput. Simul.* **127**, 114–130 (2016)
- G. Jumarie, On the representation of fractional Brownian motion as an integral with respect to $(dt)^a$. *Appl. Math. Lett.* **18**, 739–748 (2005)
- G. Jumarie, Table of some basic fractional calculus formulae derived from a modified Riemann-Liouville derivative for non-differentiable functions. *Appl. Math. Lett.* **22**(3), 378–385 (2009)
- A. Le Méhauté, *Fractal Geometry: Theory and Applications* (CRC Press, Boca Raton, 1991)
- J. Li, M. Ostoja-Starzewski, Fractal materials, beams and fracture mechanics. *ZAMP* **60**, 1–12 (2009a)
- J. Li, M. Ostoja-Starzewski, Fractal solids, product measures and fractional wave equations. *Proc. R. Soc. A* **465**, 2521–2536 (2009b); Errata (2010)
- J. Li, M. Ostoja-Starzewski, Fractal solids, product measures and continuum mechanics, chapter 33, in *Mechanics of Generalized Continua: One Hundred Years After the Cosserats*, ed. by G.A. Maugin, A.V. Metrikine (Springer, New York, 2010), pp. 315–323
- J. Li, M. Ostoja-Starzewski, Micropolar continuum mechanics of fractal media. *Int. J. Eng. Sci.* (A.C. Eringen Spec. Issue) **49**, 1302–1310 (2011)
- J. Li, M. Ostoja-Starzewski, Edges of Saturn's rings are fractal. *SpringerPlus* **4**, 158 (2015). arXiv:1207.0155 (2012)
- B.B. Mandelbrot, *The Fractal Geometry of Nature* (W.H. Freeman & Co, New York, 1982)
- G.A. Maugin, *The Thermomechanics of Nonlinear Irreversible Behaviours* (World Scientific Pub. Co., Singapore, 1999)
- G.A. Maugin, *Non-classical Continuum Mechanics: A Dictionary* (Springer, Singapore, 2016)
- W. Nowacki, *Theory of Asymmetric Elasticity* (Pergamon Press/PWN – Polish Sci. Publ., Oxford/Warszawa, 1986)
- K.B. Oldham, J. Spanier, *The Fractional Calculus* (Academic Press, San Diego, 1974)
- M. Ostoja-Starzewski, Fracture of brittle micro-beams. *ASME J. Appl. Mech.* **71**, 424–427 (2004)
- M. Ostoja-Starzewski, Towards thermomechanics of fractal media. *ZAMP* **58**(6), 1085–1096 (2007)
- M. Ostoja-Starzewski, On turbulence in fractal porous media. *ZAMP* **59**(6), 1111–1117 (2008a)
- M. Ostoja-Starzewski, *Microstructural Randomness and Scaling in Mechanics of Materials* (CRC Press, Boca Raton, 2008b)
- M. Ostoja-Starzewski, Extremum and variational principles for elastic and inelastic media with fractal geometries. *Acta Mech.* **205**, 161–170 (2009)
- M. Ostoja-Starzewski, Electromagnetism on anisotropic fractal media. *ZAMP* **64**(2), 381–390 (2013)

- M. Ostoja-Starzewski, J. Li, H. Joumaa, P.N. Demmie, From fractal media to continuum mechanics. *ZAMM* **94**(5), 373–401 (2014)
- M. Ostoja-Starzewski, S. Kale, P. Karimi, A. Malyarenko, B. Raghavan, S.I. Ranganathan, J. Zhang, Scaling to RVE in random media. *Adv. Appl. Mech.* **49**, 111–211 (2016)
- D. Stoyan, H. Stoyan, *Fractals, Random Shapes and Point Fields* (Wiley, Chichester, 1994)
- V.E. Tarasov, Fractional hydrodynamic equations for fractal media. *Ann. Phys.* **318**(2), 286–307 (2005a)
- V.E. Tarasov, Wave equation for fractal solid string. *Mod. Phys. Lett. B* **19**(15), 721–728 (2005b)
- V.E. Tarasov, Continuous medium model for fractal media. *Phys. Lett. A* **336**, 167–174 (2005c)
- V.E. Tarasov, *Fractional Dynamics: Applications of Fractional Calculus to Dynamics of Particles, Fields and Media* (Springer, Berlin, 2010)
- V.E. Tarasov, Anisotropic fractal media by vector calculus in non-integer dimensional space. *J. Math. Phys.* **55**, 083510-1-20 (2014)
- V.E. Tarasov, Electromagnetic waves in non-integer dimensional spaces and fractals. *Chaos, Solitons Fractals* **81**, 38–42 (2015a)
- V.E. Tarasov, Vector calculus in non-integer dimensional space and its applications to fractal media. *Commun. Nonlinear Sci. Numer. Simul.* **20**, 360–374 (2015b)
- H. Ziegler, *An Introduction to Thermomechanics* (North-Holland, Amsterdam, 1983)

Part IV
**Computational Modeling for Gradient
Plasticity in Both Temporal and Spatial Scales**



Modeling High-Speed Impact Failure of Metallic Materials: Nonlocal Approaches

26

George Z. Voyiadjis and Babür Deliktaş

Contents

Introduction	940
Nonlocal Strain Gradient Crystal Plasticity Formulation	943
Simple Shear of a Constrained Crystalline Strip with Double Slips	947
Experimental Procedure for Material Evaluation	954
Result of the First Set of Experiments	954
Result of the Second Set of Experiments	955
Result of the Third Set of Experiments	958
Numerical Simulations and Validating the Experimental Approach	963
Discussions and Conclusions	966
References	967

Abstract

Development and application of advanced, computationally intensive multiscale (macro-, meso-, and micro-mechanically) physically based models to describe physical phenomena associated with friction and wear in heterogeneous solids, particularly under high velocity impact loading conditions. Emphasis will be placed on the development of fundamental, thermodynamically consistent theories to describe *high-velocity material wear failure* processes in combinations of ductile and brittle materials for wear damage-related problems. The wear failure

G. Z. Voyiadjis (✉)

Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge, LA, USA

e-mail: voyiadjis@eng.lsu.edu

B. Deliktaş

Faculty of Engineering-Architecture, Department of Civil Engineering, Uludag University, Bursa, Turkey

e-mail: bdeliktas@uludag.edu.tr

© Springer Nature Switzerland AG 2019

G. Z. Voyiadjis (ed.), *Handbook of Nonlocal Continuum Mechanics for Materials and Structures*, https://doi.org/10.1007/978-3-319-58729-5_5

939

criterion will be based on dissipated energies due to plastic strains at elevated temperatures. Frictional coefficients will be identified for the contact surfaces based on temperature, strain rates, and roughness of the surfaces. In addition failure models for microstructural effects, such as shear bands and localized deformations, will be studied.

The computations will be carried with Abaqus Explicit as a dynamic temperature-displacement analysis. The contact between sliding against each other's surfaces is specified as surface-to-surface contact on the master-slave basis. The tangential behavior is defined as kinematic contact with finite sliding. The validation of computations utilizing the novel approach presented in this work is going to be conducted on the continuum level while comparing the obtained numerical results with the experimental results obtained in the laboratories in Metz, France. Reaction forces due to friction between the two specimens and temperature resulting from the dissipated energy during the friction experiment are going to be compared and discussed in detail. Additionally the indentation response at the macroscale, for decreasing the size of the indenter, will be used to critically assess and evaluate the length scale parameters.

Keywords

Failure in metals · Frictional coefficient · High speed impact · High velocity material wear · Micro structural effects · Multiscale modeling

Introduction

The need to improve the reliability and life of tribological components that are prone to severe contact stress during some of the engineering applications such as cutting tool for metals, gun tubes, engine exhaust valves, engine turbocharger components, rail gun environment, jet engine gear box splines and gears, and slippers on high-speed test track sleds. Severe contact stresses of such applications generate high temperature and create thermomechanical gouging and wear due to high-velocity sliding between contacting materials (Ireman et al. 2003; Ireman and Nguyen 2004; Arakawa 2014, 2017; Hernandez et al. 2015; Bayart et al. 2016). High-velocity sliding between dissimilar materials can result in thermomechanical gouging of the materials. This problem can be considered in the field of nonlocal modeling of heterogeneous media that assesses a strong coupling between rate-dependent plasticity (viscoplasticity) and rate-dependent damage (viscodamage or creep damage) under high-velocity gouging (Lodygowski et al. 2011; Voyiadjis et al. 2010).

The field of research dealing with wear modeling is still very elusive, yet there has been progress made in the understanding of many aspects of the wear mechanisms. The definition of the wear is stated in the work of Johansson and Klarbring (1993) as “the loss of material from a surface, transfer of material from

one surface to another or movement of material within a single surface.” This definition can be simplified as “damage to a solid surface generally involving progressive loss of materials due to relative motion between that surface and a coating surface or substance.” (Johnson et al. 1947; Klarbring 1986, 1990; Johansson and Klarbring 2000; Varga et al. 2013). A close look at the activity leading to wear is at the dimensional level of 10 nm where material structure can be seen as nanocrystalline. In addition to this small-scale phenomenon, local high spots (asperities) produce high stresses over short time scales of order of microseconds, whereas strain rate are of the order of 10^4 – 10^7 s⁻¹.

Another important aspect of wear problem is the time dependency and high temperature produced by energy dissipation through the contact interaction. Due to both high strain rate and temperature, it is possible to observe in the wear of metals the formation of adiabatic shear bands at the microscale. Many of the features described here make it obvious that the highly complex nature of the wear problem cannot be solely treated by using macroscale phenomenological models. Therefore, a clear need exists for the development of a realistic and reliable physically based material model within the framework of multiscale modeling that can be utilized in the sever contact stress applications.

The major consideration of this research is to develop an experimental/theoretical model for the material in order to better characterize and predict the internal failure surrounding the gouging and wear events. This research is to be carried out by first investigating the phenomenon of the wear, and later it will be extended to incorporate gauging problems.

The theoretical model is based on a nonlocal theory of crystal plasticity that incorporates macroscale interstate variables and their higher-order gradients in order to describe the change in the internal structure and investigate the size effect of statistical inhomogeneity of the evolution-related plasticity and damage-hardening variables. The gradients are introduced here as the hardening internal state variables and are considered independent of their local counterparts. It also incorporates the thermomechanical coupling effects as well as the internal dissipative effects through the rate-type covariance constitutive structure with a finite set of internal state variables.

Crystal plasticity models (Asaro 1983) that account for the crystalline microstructure by distinguishing between different crystallographic orientations have become successful in modeling the anisotropic plastic deformation of single crystals. The discrete slip systems that depend on the crystal structure are incorporated into these models, but the plastic part of deformation is still modeled in a continuum sense. The characteristic length of the deformed microstructure becomes significant in the analysis of the material at a scale where the microstructure characteristic length is no longer negligible with respect to the material size. This triggers an important question, whether and how macroscopic overall mechanical properties as strength, hardness, etc. depend upon a natural internal length scale related to the characteristic size of the microstructures in the material. A nonlocal crystal plasticity (NCP) model is developed in this work using thermodynamically consistent higher-order strain gradient theory. Multiple length

scales are introduced at the NCP to account for the microstructure. These length scales are used to characterize size effects, strengthening, hardening, values of dislocation line energies, back-stress magnitudes, and the thickness of the plastic boundary layer.

The theoretical model is formulated for the crystalline system with multiple symmetric double slip systems. The boundary value problem analyzed here is the simple shear of a constrained crystalline strip in order to investigate the effects of the dissipative and energetic parameters on the size effect response of the constrained crystalline strip systems. The developed theory described quantitatively the thickness of the boundary layer, hardening, and strengthening response of the system.

The classical crystal plasticity, which inherently includes no material length scales, cannot predict size effects. Strain gradient plasticity (SGP) theories extend the classical plasticity models by including an intrinsic material length scale and are therefore appropriate for problems involving small dimensions (Aifantis and Willis 2006; Anand et al. 2005; Fleck and Willis 2008; Fredriksson and Gudmundson 2007). These material length scales are necessary both for dimensional consistency when strain gradients are used in the formulation and in crystal plasticity based on a continuum description of the dislocation behavior. The physical basis of the SGP for metals has been founded on theoretical developments concerning geometrically necessary dislocations (GNDs) (Nye 1953). Standard micromechanical modeling of metals for the inelastic material behavior of single crystals and polycrystals is commonly based on the premise that resistance to glide is due mainly to the random trapping of mobile dislocations during locally homogeneous deformations. Such trapped dislocations are commonly referred to as statistically stored dislocations (SSDs) and act as obstacles to stop further dislocation motion, resulting in hardening. An additional contribution to the density of immobile dislocations and to hardening can arise when the continuum length scale approaches that of the dominant microstructural features.

One of the open issues of the SGP is the ongoing discussion of the energetic and dissipative nature of the dislocation network that accounts for many plasticity phenomena. Gurtin (2003) argued that the density of geometrically necessary dislocations is quantified by Nye's tensor that leads to increase in the free energy. However, Fleck and Willis (2008) discussed this issue by questioning whether the additional strengthening is mainly energetic or dissipative. They assumed that the core energy of dislocations stored during plastic deformation is much smaller than the plastic work dissipated in dislocation motion. Based on this observation, they concluded that both statistically stored and geometrically necessary dislocations contribute more to plastic dissipation than to a change in energy. Bardella (2006, 2007) also pointed out that modeling involving only energetic material length scales may not be sufficient to describe the size effects exhibited in metals. He reasoned the fact that energetic length scales, defined through a function of Nye's dislocation density tensor, allow the description of the increase in strain hardening accompanied with diminishing size, but they do not help in capturing the related strengthening. He

proposed that at least one dissipative length scale is required in modeling in order to capture the strengthening effect.

In this work both energetic and dissipative material length scales are introduced, and their physically based relations as a function of the accumulated plastic deformation, strain rate, temperature, and corresponding microstructural features are proposed. It was observed that plasticity and damage phenomena at small-scale levels dictate the necessity of more than one length scale parameters in the gradient description. Energetic material length scales quantify the size effects due to GNDs. The need of more energetic material length scales in order to have a better description of the material behavior should be first evaluated in light of the experimental results (or discrete dislocation simulations). Two main energy storage mechanisms are used here. The first is that energy stored in dislocation network uses dislocation density that depends on the inelastic strain path. Dislocation density increases with inelastic deformation and eventually reaches a saturation point. The second storage mechanism is due to the presence of the second hard phase.

Nonlocal Strain Gradient Crystal Plasticity Formulation

In this section the nonlocal strain gradient crystal plasticity theory is presented. This theory is developed based on the thermodynamically consistent higher-order strain gradient theory developed by Voyiadjis and coworkers (Voyiadjis and Deliktas 2009a, b; Voyiadjis and Abu Al-Rub 2007). In order to adapt this theory to a crystalline system, they introduced in the virtual work the following fields $X^{(\beta)}$ and $\mathbb{X}^{(\beta)}$ that are work-conjugates to the slip, $\gamma^{(\beta)}$, and the slip gradient, $\gamma_{,i}^{(\beta)}$, respectively. The principle of virtual work can be expressed in terms of these fields as follows:

$$\int_V \left(\sigma_{ij} \delta \varepsilon_{ij}^e + \sum_{\beta} X^{(\beta)} \delta \gamma^{(\beta)} + \sum_{\beta} \mathbb{X}_{,i}^{(\beta)} \delta \gamma_{,i}^{(\beta)} \right) dV = \int_S t_i \delta u_i dS + \int_S \mathfrak{t}^{(\beta)} \delta \gamma^{(\beta)} dS \quad (1)$$

where σ_{ij} is the standard stress tensor with $\sigma_{ij} = \sigma_{ji}$, t_i is the standard traction stress vector, while $\mathfrak{t}^{(\beta)}$ is the nonstandard traction force work-conjugate to the slip, $\gamma^{(\beta)}$. In order to derive balance laws, integration by parts is used on some terms of the integral in Eq. (1) along with the divergence theorem ($\int_V \text{grad} (F) dV = \int_V \nabla \cdot \mathbf{F} dV = \int_S F \mathbf{n} dS$). The gradient of the displacement is given by $\varepsilon_{ij} = 1/2(u_{i,j} + u_{j,i})$, and the resolved shear stress is $\tau^{(\beta)} = m_i^{(\beta)} \sigma_{ij} s_j^{(\beta)}$ where $m_i^{(\beta)}$ and $s_j^{(\beta)}$ are unit vectors characterizing the normal to the slip plane β and its slip direction, respectively. Equation (1) is now expressed as follows:

$$\int_V \left(\sigma_{ij,j} \delta u_i + \sum_{\beta} \left(X^{(\beta)} - \tau^{(\beta)} - \mathbb{X}_{i,i}^{(\beta)} \right) \delta \gamma^{(\beta)} \right) dV + \int_S (t_i - \sigma_{ij} n_j) \delta u_i dS + \int_S \left(\mathfrak{t}^{(\beta)} - \mathbb{X}_i^{(\beta)} n_i \right) \delta \gamma^{(\beta)} dS \tag{2}$$

Equation (2) must hold for any variations in δu_i and $\delta \gamma^{(\beta)}$, and therefore one obtains the classical equilibrium equation

$$\sigma_{ij,j} = 0 \tag{3}$$

and the micro-force balance equation

$$X^{(\beta)} - \tau^{(\beta)} - \mathbb{X}_{i,i}^{(\beta)} = 0 \tag{4}$$

Standard boundary conditions are now obtained that can be defined by prescribing either traction force or displacement, respectively, at each point on the boundary

$$t_i = \sigma_{ij} n_j \quad \text{or} \quad \bar{u}_i \tag{5}$$

Nonstandard boundary conditions can be given by prescribing either micro-traction force or slip at each point of the boundary

$$\mathfrak{t}^{(\beta)} = \mathbb{X}_i^{(\beta)} n_i \quad \text{or} \quad \bar{\gamma}^{(\beta)} \tag{6}$$

A boundary condition of $\bar{\gamma}^{(\beta)} = 0$ form is referred to as a micro-clamped boundary condition, and a boundary condition of $\mathfrak{t}^{(\beta)} = 0$ is termed a micro-free boundary condition.

In fact, the plastic strain, ε_{ij}^p , in a single crystal can be expressed as the symmetric part of the plastic distortion, γ , by the following expression:

$$\gamma_{ij} = \sum_{\beta=1}^N \gamma^{(\beta)} m_i^{(\beta)} s_j^{(\beta)} \tag{7}$$

Therefore, the plastic deformation can be expressed in terms of the plastic distortion as

$$\varepsilon_{ij}^p = \sum_{\beta=1}^N \gamma^{(\beta)} \mu_{ij}^{(\beta)} \tag{8}$$

where

$$\mu_{ij}^{(\beta)} = \frac{1}{2} \left(m_i^{(\beta)} s_j^{(\beta)} + m_j^{(\beta)} s_i^{(\beta)} \right) \quad (9)$$

and $\mu_{ij}^{(\beta)}$ is the symmetric Schmid tensor of the slip system β .

A motivation for the nonlocal formulation proposed by Gurtin (2003) stems from the fact that plastic deformation in crystalline solids arises from the motion of dislocations, which are line defects characterized by the Burgers vector b_i . This observation allows one to characterize the net Burgers vector B_i with respect to a closed loop C in the crystal through the following integral

$$B_i = \oint_C u_{ij}^p dx_j = \oint_S \alpha_{ij} v_j dS \quad (10)$$

in which S is any surface with boundary C and v_j is the unit normal field for S suitably oriented with respect to C . The tensor α_{ij} , which is often referred to as Nye's dislocation density tensor (Nye 1953) or as the density of geometrically necessary dislocations, is given by the explicit relation

$$\alpha_{ij} = \sum_{\beta=1}^N e_{jkl} \gamma_{,k}^{(\beta)} m_i^{(\beta)} s_l^{(\beta)} \quad (11)$$

where e_{jkl} is the alternating tensor.

Helmholtz free energy that accounts only for stored energy allows one to obtain energetic components of the thermodynamic forces σ_{ij} , $X^{(\beta)}$, and $\mathbb{X}^{(\beta)}$ work-conjugates to the elastic strain, ε_{ij}^e , slip, $\gamma^{(\beta)}$, and slip gradient, $\gamma_{,i}^{(\beta)}$, respectively. The functional form of the Helmholtz free energy is assumed to be

$$\Psi = \frac{1}{2} \varepsilon_{ij}^e E_{ijkl} \varepsilon_{kl}^e + \sum_{\beta=1}^N \frac{1}{2} a_1^{(\beta)} \left(\gamma^{(\beta)} \right)^2 + \frac{1}{2} a_2 \alpha_{ij} \alpha_{ij} \quad (12)$$

In this formulation one does not assume that Nye's tensor α_{ij} is an independent state variable of the Helmholtz free energy but it is dependent on the slip gradient, $\gamma_{,i}$, of the free energy. The energetic components of the thermodynamic forces can be obtained as follows:

$$\begin{aligned} \sigma_{ij} &= \frac{\partial \Psi}{\partial \varepsilon_{kl}^e} = E_{ijkl} \varepsilon_{kl}^e \\ {}^{en} X^{(\beta)} &= \frac{\partial \Psi}{\partial \gamma^{(\beta)}} = a_1^{(\beta)} \gamma^{(\beta)} \\ {}^{en} \mathbb{X}_i^{(\beta)} &= \frac{\partial \Psi}{\partial \gamma_{,i}^{(\beta)}} = \frac{\partial \Psi}{\partial \alpha_{kl}} \frac{\partial \alpha_{kl}}{\partial \gamma_{,i}^{(\beta)}} = a_2 \alpha_{kl} N_{kil}^{(\beta)} \end{aligned} \quad (13)$$

where $N_{kjl} = e_{ijl}m_i^{(\beta)}s_k^{(\beta)}$, $a_1^{(\beta)} = h_o$ and $a_2 = G\ell_{en}^2$. The thermodynamic force given by (13)₃ is similar to the expression for $\xi_i^{(\beta)}$ that is defined by Bittencourt et al. (2003). ℓ_{en} is the energetic material length scale that quantifies the size effect due to GND (related to the Burgers vector length). This characteristic length scale from a microstructural point of view can be attributed to dislocation network such as dislocation spacing, dislocation cell size, and misorientation of the cells in the metallic materials.

In order to satisfy the second law of thermodynamics, one requires the dissipation to be positive. This implies that $\sum_{\beta} X^{(\beta)}\dot{\gamma}^{(\beta)} + \sum_i X_i^{(\beta)}\dot{\gamma}_{,i}^{(\beta)} \geq 0$. This inequality leads to the existence of the dissipation potential that should be a convex functional form in terms of its variables. One possible choice of such potential can be given by the following power law (Bardella 2007)

$$D = \frac{h_o\gamma_o}{N + 1} \left(\frac{\gamma_{eff}^{(\beta)}}{\gamma_o} \right)^{N+1} \tag{14}$$

where $\gamma_{eff}^{(\beta)} = \sqrt{|\gamma^{(\beta)}|^2 + \ell_{dis}^2\gamma_i^{(\beta)}\gamma_{,i}^{(\beta)}}$. ℓ_{dis} is the dissipative material length scale that quantifies the strengthening due to the size effect. This characteristic length scale from a microstructural point of view can be attributed to the movement of the mobile dislocations.

The dissipative components of each thermodynamic force can be obtained from Eq. (14) by taking partial derivatives of this potential with respect to its flux respective variables

$$\begin{aligned} dis X^{(\beta)} &= h_o \left(\frac{\gamma_{eff}}{\gamma_o} \right)^{N-1} \frac{\gamma^{(\beta)}}{\gamma_o} \\ dis X_i^{(\beta)} &= h_o \ell_{dis}^2 \left(\frac{\gamma_{eff}}{\gamma_o} \right)^{N-1} \frac{\gamma_{,k}^{(\beta)}}{\gamma_o} \end{aligned} \tag{15}$$

Making use of Eqs. (13)₂, (13)₃, and (15) into (4) and (6), the following nonlocal differential equation for the flow stress and the nonstandard boundary condition can be obtained, respectively,

$$\begin{aligned} \tau^{(\beta)} &= a_1^{(\beta)}\gamma^{(\beta)} + h_o \left(\frac{\gamma_{eff}}{\gamma_o} \right)^{N-1} \frac{\gamma^{(\beta)}}{\gamma_o} \\ &\quad - \ell_{en}^2 G \left[\alpha_{ij} N_{ikj}^{(\beta)} \right]_{,k} - h_o \ell_{dis}^2 \left[\left(\frac{\gamma_{eff}}{\gamma_o} \right)^{N-1} \frac{\gamma_{,k}^{(\beta)}}{\gamma_o} \right]_{,k} \end{aligned} \tag{16}$$

and for the corresponding nonstandard boundary condition

$$\mathfrak{t}^{(\beta)} = \ell_{en}^2 G \alpha_{ij} N_{ikj}^{(\beta)} + h_o \ell_{dis}^2 \left[\left(\frac{\gamma_{eff}}{\gamma_o} \right)^{N-1} \frac{\gamma_{,k}^{(\beta)}}{\gamma_o} \right] n_i \tag{17}$$

The elasticity is taken to be isotropic, characterized by Young’s modulus E and Poisson’s ratio ν , with the shear modulus given by $2G = E/(1 + \nu)$. Thus there are seven positive material parameters that characterize the mechanical behavior: E , ν , h_o , Γ_o , N and material length scales, ℓ_{dis} and ℓ_{en} , respectively. In addition, for each crystal the number of slip systems and the respective orientation of each system specified by $m_i^{(\beta)}$, $s_i^{(\beta)}$ ($m_i^{(\beta)}s_i^{(\beta)} = 0$) need to be identified.

Simple Shear of a Constrained Crystalline Strip with Double Slips

The boundary value problem analyzed here is the simple shear of a constrained crystalline strip, of height H in the x_2 direction and unbounded in the x_1 and x_3 directions. The crystal is characterized by incompressible isotropic linear elastic material with plain strain and quasi-static loading conditions. It is also assumed that the crystalline strip has multiple double β slip systems that consist of two possible glides symmetrically oriented with respect to any plane constant x_2 by angle θ_β . The strip is sheared by applying to the plane $x_2 = H$ a uniform displacement equal to $Y(t)H$ in the x_1 direction where $Y(t)$ is the prescribed shear strain. The schematic description of this problem is illustrated in Fig. 1.

The standard *macroscopic boundary conditions* are

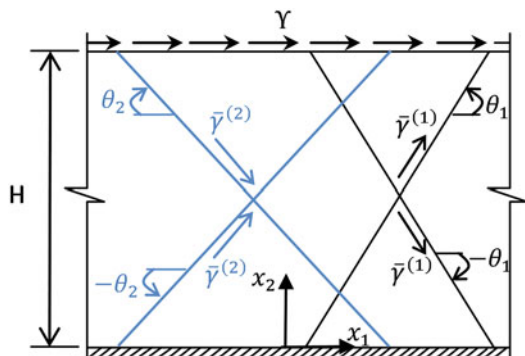
$$\begin{aligned} u_1 &= 0, \text{ along } x_2 = 0 \\ u_1 &= U(t) = HY(t), \quad u_2 = 0 \text{ along } x_2 = H \end{aligned} \tag{18}$$

In the constrained layer problem, one restricts to monotonic loading, so that the prescribed shear rate satisfies $\dot{Y} > 0$.

The nonstandard boundary conditions due to the nonlocality are defined as *micro-free boundary conditions on the sides* of the strip

$$\mathfrak{t}^{(\beta)} = \mathfrak{X}_i^{(\beta)} n_i = 0 \text{ along } x_1 = \pm w \tag{19}$$

Fig. 1 A constrained crystalline strip with multiple symmetric double slip systems subjected to the simple shear



and *micro-clamped boundary conditions* on the bottom and top of the strip

$$\gamma^{(\beta)} = 0, \text{ along } x_2 = 0, H \tag{20}$$

One now considers the solution of this boundary value problem with field quantities that are independent of x_1 . Macroscopic equilibrium requires that σ_{12} be spatially uniform. One seeks the solution for which σ_{11} and σ_{22} vanish. The elasticity solution in simple shear has a spatially uniform stress. Thus the entire strip reaches yield at the same time. The orientation of the first slip of the double system is given by

$$m_i^1 = \cos \theta_\beta \mathbf{e}_1 + \sin \theta_\beta \mathbf{e}_2, \quad s_i^1 = -\sin \theta_\beta \mathbf{e}_1 + \cos \theta_\beta \mathbf{e}_2 \tag{21}$$

The orientation of the second slip of the double system can be obtained by replacing θ_β by $-\theta_\beta$ and recalling that the following equalities hold due to symmetry ($\gamma_1^{(\beta)} = \gamma_2^{(\beta)} = \bar{\gamma}^{(\beta)}$) and ($\tau_1^{(\beta)} = \tau_2^{(\beta)} = \bar{\tau}^{(\beta)}$). The sole nonvanishing component of the strain tensor is given by

$$\varepsilon_{12}^p = \sum_{\beta} \bar{\gamma}^{(\beta)} \cos 2\theta_\beta \tag{22}$$

Since all the fields must be independent of x_1 and x_3 and due to symmetry, therefore the nonvanishing component of the Nye’s tensor is in the x_2 direction due to the edge dislocation lying along x_3 and is given by

$$\alpha_{23} = \sum_{\beta} \bar{\gamma}_{,2}^{(\beta)} \sin^2 \theta_\beta \tag{23}$$

Hence the only nonvanishing components of N_{ijk} tensor can be obtained as

$$N_{223} = \frac{\partial \alpha_{23}}{\partial \gamma_{,2}^{(\beta)}} = \sin^2 \theta_\beta \tag{24}$$

One first considers the linear case by setting $N = 1$ in Eqs. (16) and (17) and making use of Eqs. (21–24); the simplified form of the Eqs. (16) and (17) can be obtained as follows:

$$\bar{\tau}^{(\beta)} = h_o \bar{\gamma}^{(\beta)} + h_o \frac{\bar{\gamma}^{(\beta)}}{\gamma_o} - \ell_{en}^2 G \sin^2 \theta_\beta \sum_{\zeta} \bar{\gamma}_{,22}^{\zeta} \sin^2 \theta_{\zeta} - h_o \ell_{dis}^2 \frac{\bar{\gamma}_{,22}^{(\beta)}}{\gamma_o} \tag{25}$$

and for the corresponding nonstandard boundary condition

$$\mathfrak{t}^{(\beta)} = \left[\ell_{en}^2 G \sin^2 \theta_\beta \sum_{\zeta} \bar{\gamma}_{,2}^{\zeta} \sin^2 \theta_{\zeta} + h_o \ell_{dis}^2 \frac{\bar{\gamma}_{,2}^{(\beta)}}{\gamma_o} \right] n_2 \tag{26}$$

Equation (25) can be written in the matrix form as

$$\mathcal{M}^{(\beta\zeta)}\gamma_{22}^{(\zeta)} - \gamma^{(\beta)} = -\mathcal{F}^{(\beta)} \quad (27)$$

where $\mathcal{M}^{(\beta\zeta)}$ is defined as

$$\mathcal{M}^{(\beta\zeta)} = \begin{bmatrix} a_1^2 + b_1^2 & b_1 b_2 & \cdots & \cdots & b_1 b_\zeta \\ b_2 b_1 & a_2^2 + b_2^2 & & & \\ \vdots & & \ddots & & \\ \vdots & & & \ddots & \\ b_\beta b_1 & & & & a_\beta^2 + b_\beta^2 \end{bmatrix} \quad (28)$$

and its components are defined as

$$\begin{aligned} a_\beta &= \frac{\ell_{dis}}{\sqrt{1+\gamma_o}}, \quad b_\beta = \sqrt{\frac{G\gamma_o}{(1+\gamma_o)h_o}} \ell_{en} \sin^2 \theta_\beta \\ \mathcal{F}^\beta &= \frac{\tau^{(\beta)}\gamma_o}{(1+\gamma_o)h_o} \end{aligned} \quad (29)$$

For the sake of simplicity, the strip system is assumed to have only one double system, and the dissipative material scale is set to zero. Equation (27) reduces to the following one dimensional form

$$\bar{\gamma}_{,22} - \lambda^2 \bar{\gamma} = -\lambda^2 \quad (30)$$

where the coefficients λ^2 and F are defined as

$$\lambda^2 = \frac{(1+\gamma_o)h_o}{(\ell)^2 G \gamma_o \sin^4 \theta}, \quad F = \frac{\bar{\tau}}{(\ell)^2 G \sin^4 \theta} \quad (31)$$

Case 1 No dissipative hardening ($h_o = 0$, such that $\lambda^2 = 0$) is considered here. In this case the solution of Eq. (30) with the boundary conditions given by Eq. (20) yields the following quadratic relation for the plastic slip as a function of x_2 :

$$\bar{\gamma} = \frac{F}{2} (x_2 H - x_2^2) \quad (32)$$

The remaining field quantities ε_{12} and σ_{12} can be computed using the following steps. The nonvanishing components of the plastic strain given by Eq. (22) can be rewritten as follows:

$$\varepsilon_{12}^p = \gamma \cos 2\theta \quad (33)$$

and the resolved shear stress is given by $\bar{\tau} = \cos 2\theta \sigma_{12}$ where shear stress is defined as $\sigma_{12} = 2G (\varepsilon_{12} - \bar{\gamma} \cos 2\theta)$. Averaging shear stress with respect to x_2 over the

interval $[0, H]$ and knowing that σ_{12} is spatially constant and the average of ε_{12} over the interval $[0, H]$ is $\Upsilon/2$, then one obtains

$$\sigma_{12} = G (\Upsilon - 2 \cos 2\theta \gamma_{ave}) \tag{34}$$

The solution of the differential Eq. (30) that is expressed by Eq. (32) is a function of $\bar{\tau}$. By eliminating σ_{12} from Eq. (34), one can solve the resulting linear equation for $\bar{\tau}$ such as

$$\bar{\tau}(r) = \frac{6G\Upsilon \cos 2\theta \sin^4 2\theta}{(6\sin^4 2\theta + \cos^2 2\theta \gamma_o r^2)} \tag{35}$$

where $r = H/\ell$. Substituting Eq. (35) into (32) and (34) along with setting $(x_2 = zH)$ yields the following relation:

$$\bar{\gamma}(r) = \frac{\bar{\tau}(r) r^2}{2G \sin^4 2\theta} (z - z^2) \tag{36}$$

This solution shows that the slip distribution is in quadratic form in the normalized thickness. The shear stress component, σ_{12} , can be obtained in terms of

$$\sigma_{12}(r) = \frac{6G\Upsilon \sin^4 2\theta}{(6\sin^4 2\theta + \cos^2 2\theta \gamma_o r^2)} \tag{37}$$

These two equations are plotted in Fig. 2a, b, respectively.

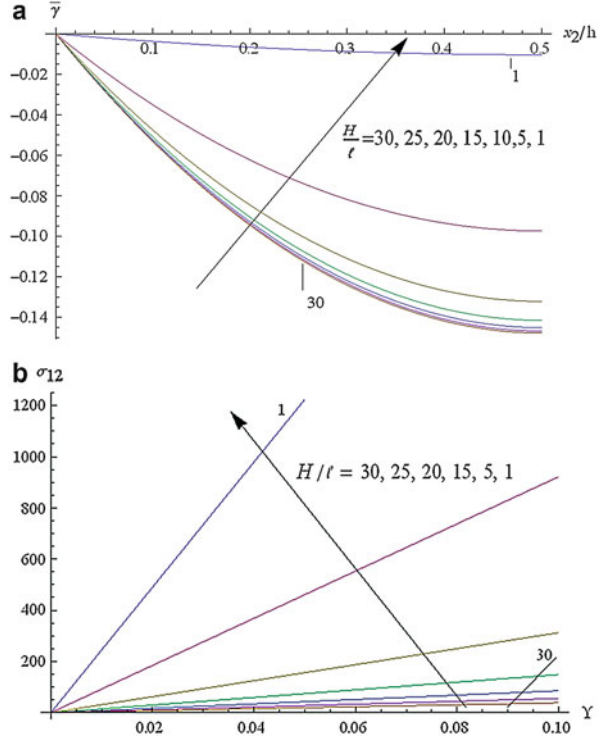
The size of the thickness is defined by $(r = H/\ell)$. Since ℓ is the material property and has a constant value, the value of r represents the size of the thickness. It is clear from both figures that the material indicates strengthening with decrease in size. Figure 2a, b illustrate this fact by showing the decrease in thickness results in an increase in shear stress which in turn provides less plastic deformation. Another observation from these analyses is the formation of the boundary layer with decreasing thickness (region of gradients). It is obvious from Fig. 2a that the boundary layer starts forming after r reduces to 20. It is clearer when r is 15.

Case 2 In the case of the dissipative hardening where $(h_o > 0)$, the solution of the differential equation in Eq. (30) along with the boundary conditions given in Eq. (20) results to the following expression:

$$\bar{\gamma} = \frac{F}{\lambda^2} [1 - \cosh [\lambda x_2] + \sinh [\lambda x_2] \tanh [H\lambda/2]] \tag{38}$$

By averaging the Eq. (38) over the interval $[0, H]$ and substituting the resulting relation into Eq. (34), one can solve the resulting equality for $\bar{\tau}$

Fig. 2 Ideal plastic case where ($h_o = 0$) at various length scales ($H/\ell = 1, 5, 10, 15, 20, 25, 30$). Crystal strip has two slip systems with ($\theta^1 = 60^\circ, \theta^2 = -60^\circ$), $G = 26300$ MPa, $\gamma_o = 0.01$, $\ell_{dis} = 0$, and $N = 1$. (a) Shear slip distribution vs. normalized thickness. (b) Shear stress distribution vs. imposed shear strain



$$\bar{\tau} = \frac{G\Upsilon}{\left(\frac{1}{\cos 2\theta} + \frac{2G\gamma_o}{h_o(\gamma_o+1)} \cos 2\theta \left\{ 1 - \frac{\tanh[H\lambda/2]}{H\lambda/2} \right\} \right)} \quad (39)$$

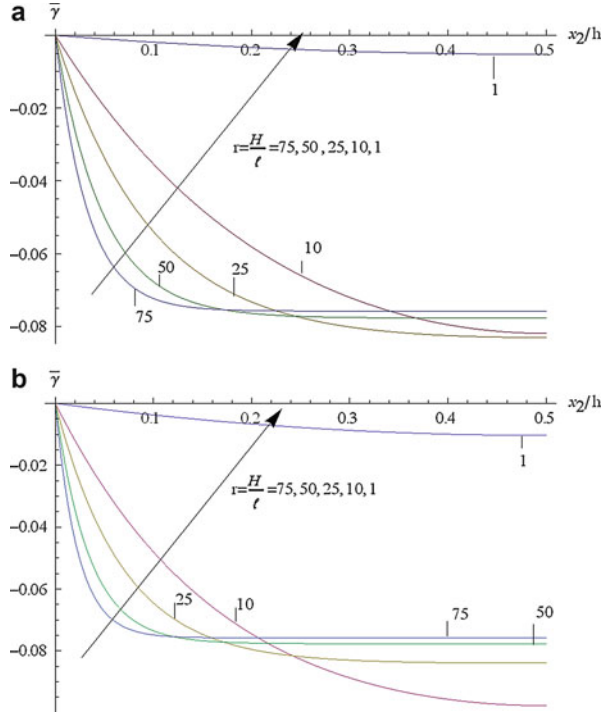
Substituting Eq. (39) into (32) and (34) yields the following relations. Shear stress component, σ_{12} , can be obtained as a function of $r = H/\ell$ by replacing the term $H\lambda/2$ with Ar where $A = \sqrt{h_o(\gamma_o + 1)}/G/2\sin^2\theta$

$$\sigma_{12}(r) = \frac{G\Upsilon}{\left[1 + \frac{2G\gamma_o}{h_o(\gamma_o+1)} \cos^2 2\theta \left\{ 1 - \frac{\tanh[Ar]}{Ar} \right\} \right]} \quad (40)$$

In the case when $r \rightarrow 0$, the response of the shear stress is $\sigma_{12}(r) = G\Upsilon$ which is the linear elastic solution. For the case when $r \rightarrow \infty$, the response of the shear stress is obtained as $\sigma_{12}(r) = G\Upsilon/(1 + 2G\gamma_o \cos 2\theta/h_o(\gamma_o + 1))$ which shows size-dependent behavior. Figure 3 shows the curves for the plastic strain versus normalized thickness for various length scales.

Although the predictions of the models show similar trends, there is variation in their hardening behaviors. This is because the proposed model introduces decomposition for each thermodynamics force. However, in the works of Bardella

Fig. 3 Comparison of two models of the plastic slip distribution vs. normalized thickness at various length scales ($H/\ell = 1, 10, 25, 50, 75$) that are obtained from the (a) proposed theory and (b) formulation by Bardella (2006). Crystal strip has two slip system with ($\theta^1 = 60^\circ, \theta^2 = -60^\circ$), $G = 26300$ MPa, $h_o = 50$ MPa, $\gamma_o = 0.01$, $Y = 0.1, \ell_{dis} = 0$, and $N = 1$



(2007) and Anand et al. (2005), they decomposed the thermodynamic forces that are related to the gradient. The current model indicates that the hardening becomes effective when the length scale parameter is reduced to 25, whereas the prediction of the other model (Bardella 2007) shows hardening at a value of 10. This explains that the current model predicts the thickness of the boundary layer higher than the prediction of that by Bardella (2007). Figure 4 shows the predictions of both models on the response of the shear stress versus imposed shear strain.

If one looks at the marked dashed lines in both Fig. 4a, b, it is clear that both models agree in their predictions of strengthening due to size effects.

A parametric study is performed by plotting the shear stress versus the slip orientation. This result is presented in Fig. 5.

As observed from Fig. 5, the influence of the orientation of the slip planes with variation in θ indicates clearly the length scale effects. In the case when the orientation of the slips is $\theta \cong 0^\circ$, the strip becomes less stiff, and the most compliant response is observed. However, when the orientation of the slips is around $\theta \cong 45^\circ$, one obtains a stiffer response, and the maximum shear stress is obtained. In both cases since the material response is linear elastic, no size effect is observed.

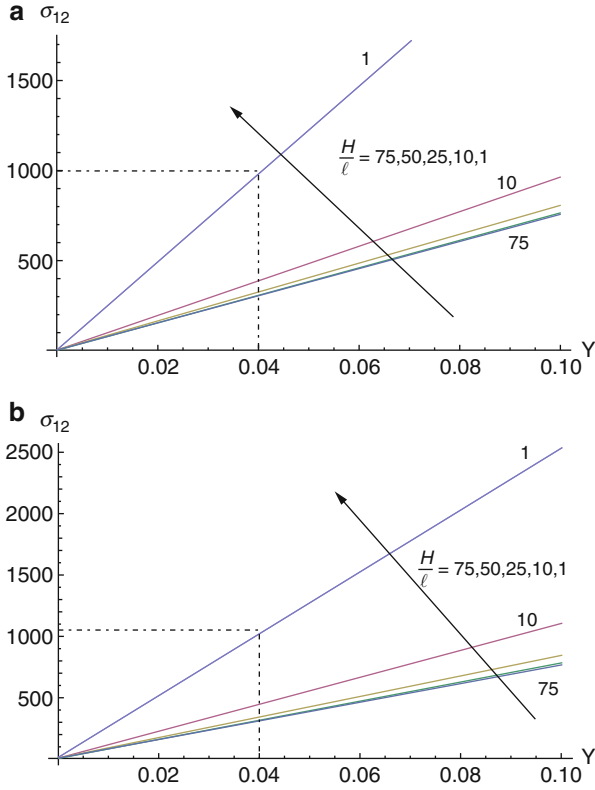


Fig. 4 Comparison between two models for the shear stress vs. imposed shear strain at various length scales ($H/l = 1, 10, 25, 50, 75$) that are obtained from the (a) proposed theory and (b) formulation by Bardella (2006). Crystal strip has two slip systems with ($\theta^1 = 60, \theta^2 = -60$), $G = 26300$ MPa, $h_o = 50$ MPa, $\gamma_o = 0.01$, $\ell_{dis} = 0$, and $N = 1$

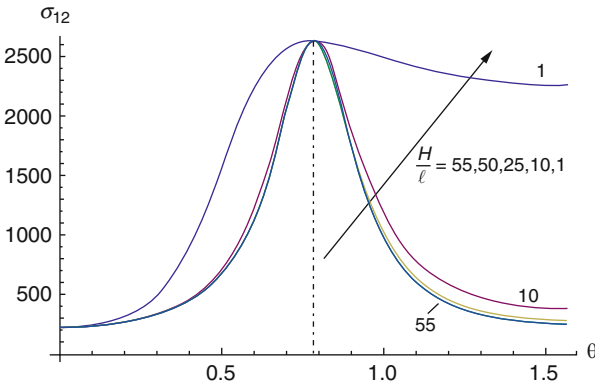


Fig. 5 Shear stress distribution versus orientation of the two symmetric slip systems at various length scales ($H/l = 1, 10, 25, 50, 55$). Other parameters used are $G = 26300$ MPa, $h_o = 50$ MPa, $\gamma_o = 0.01$, $Y = 0.1$, $\ell_{dis} = 0$, and $N = 1$. Angles are measured in radians

Table 1 Specimen designations and measured actual cross-sectional dimensions

Specimen designation	Specimen material	Section dimensions (in × in)	Applied loading
1080_04_1	AISI 1080	0.039 × 0.4	Monotonic increasing
1080_04_2	AISI 1080	0.039 × 0.4	Cyclic
4130N_05_1	AISI 4130	0.056 × 0.315	Monotonic increasing
4130N_05_2	AISI 4130	0.056 × 0.315	Cyclic
4130N_05_3	AISI 4130	0.056 × 0.316	Cyclic
VAX300_06_1	VascoMax C-300	0.064 × 0.313	Monotonic increasing
VAX300_06_2	VascoMax C-300	0.065 × 0.316	Cyclic
VAX300_06_3	VascoMax C-300	0.064 × 0.314	Cyclic

Experimental Procedure for Material Evaluation

The experimental procedure consists of three sets of tests conducted on AISI 1080, AISI 4130, and VascoMax C-300 steels. The first set of experiments involve monotonic tensile loading of the specimens up to failure, while they were unloaded at certain levels and reloaded in the second and the third sets of experiments.

An MTS 810 Hydraulic Materials Testing System was used to apply uniaxial stresses on the specimens. The strain data were obtained using a 1 in. gauge length MTS extensometer (Model No. 634.11E–24) and 1 mm gauge length Omega general-purpose pre-wired 1-axis strain gauges. The actual cross-sectional dimensions of the specimens were measured prior to tests using a digital caliper, and the measured dimensions are presented in Table 1 with the corresponding designations. The three sets of experiments and their results are presented next.

Result of the First Set of Experiments

In the first phase of the experimental program, one specimen from each material type was tested under monotonic increasing tensile loads in order to obtain mechanical properties (yield strength and ultimate strength) of the materials. During the tests, only the longitudinal strain values were recorded in the midsection of the specimens with a 1 in. gauge length MTS extensometer (Fig. 6).

The rate of loading of the MTS testing machine was set to be 0.01 in./min for specimen 4130N_05_1; however, due to the large size of the output file, it was decreased to 0.02 in./min for specimens 1080_04_1 and Vax300_06_01. Determinations of these loading rate values were made based on ASTM E8, Standard Test Methods for Tension Testing of Metallic Materials. Even though the recommended rates of loading values in ASTM E8 are higher than the ones used in this study, slower loading rates were used in this study because the cross sections of

Fig. 6 Measurement of longitudinal strains



the specimens are considerably smaller than the standard tension testing specimens addressed in the standard.

Using the output data file obtained from the software controlling the MTS testing machine, the engineering stress values were obtained by dividing the actual stress output data by the measured initial cross-sectional areas, and the stress-strain curves for all three tested materials are presented in Figs. 7, 8, and 9.

Result of the Second Set of Experiments

For the second set of experiments, transverse strain values were also recorded in addition to the longitudinal strains. For this purpose, a pair of pre-wired Omega strain gauges was attached on opposite surfaces of one specimen from each material type (Fig. 10). Due to the small width of the specimens, 1 mm gauge length (total length of 4 mm) strain gauges were used. The strain gauges were placed at the midsections of the specimens using commercial strain-gauge glue. Through a data logger, the strain data from both of the strain gauges are stored in micro-strains.

The second phase of the experimental program differs from the first one not only for measuring transverse strains but also in terms of the loading procedure. The specimens, in the second experimental phase, were loaded up to certain stress/strain levels, and then they were unloaded until the applied force is equal to zero.

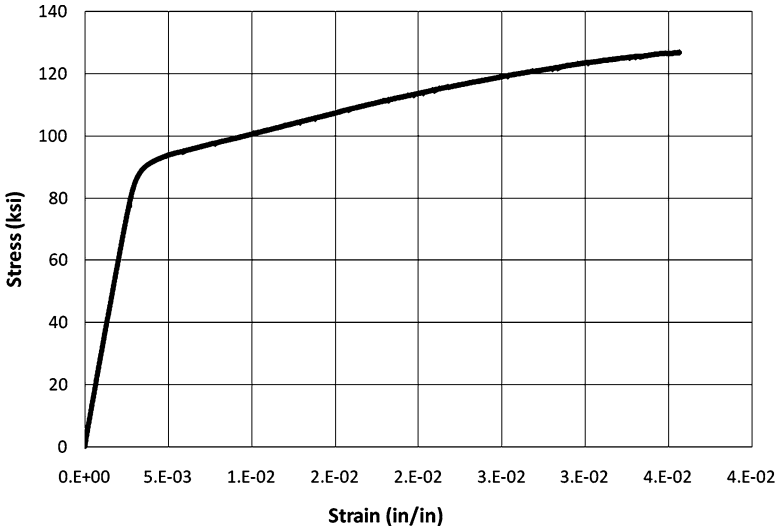


Fig. 7 Calculated stress-strain curve of specimen 1080_04_1 under monotonic loading

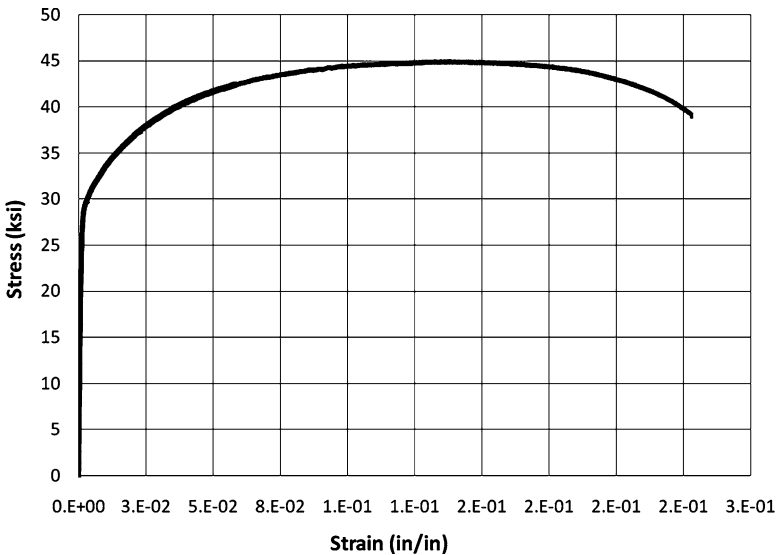


Fig. 8 Calculated stress-strain curve of specimen 4130N_05_1 under monotonic loading

Several cycles were decided to be applied in such a way that the peak strain value of each cycle exceeds the one that belongs to the previous cycle. For this purpose, special test procedures were defined using the MTS controller software, and the time, machine displacement, applied force, and strain values were stored.

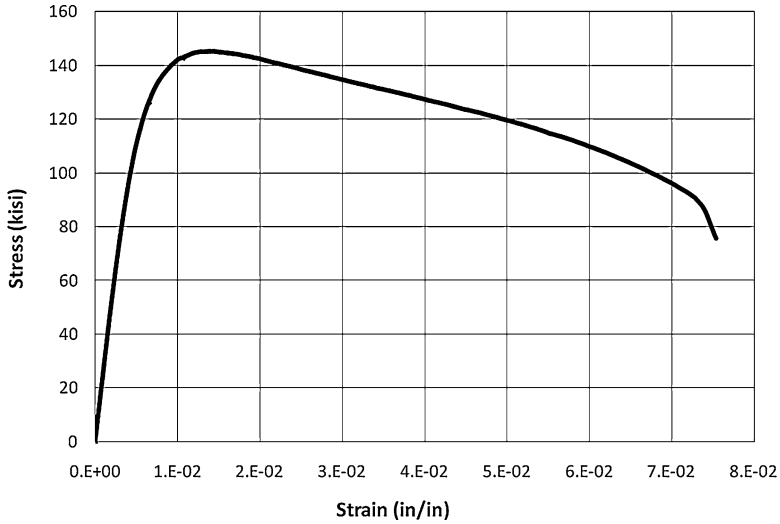
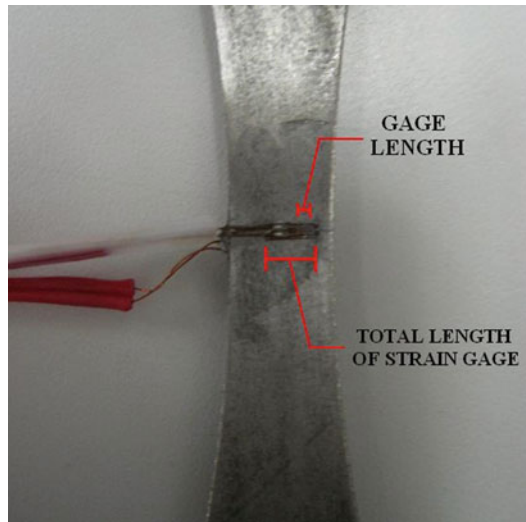


Fig. 9 Calculated stress-strain curve of specimen VAX300_06_1 under monotonic loading

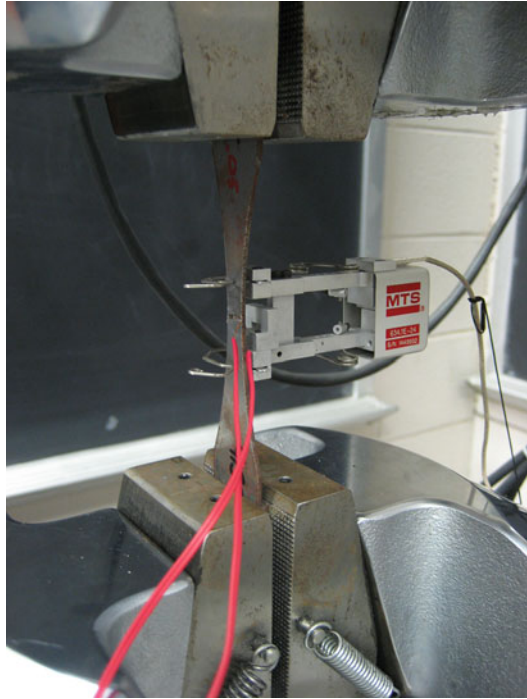
Fig. 10 A pair of pre-wired strain gauges glued on opposite sides of a specimen



In order to calculate the applied stresses, the average of the transverse strain obtained from the two strain gauges was used. The transverse strain along the thickness of the specimen was assumed to be equal to the transverse strain along the width. Figure 11 illustrates how the extensometer and the strain gauges were used to measure longitudinal and transverse strains.

The transverse strain data obtained from the attached strain gauges were averaged to be used for calculating the change in the cross section of the specimens, thus the applied stresses. The strain data collected from the strain gauges and their average

Fig. 11 Measurement of longitudinal and transverse strains



values are plotted and shown in Figs. 8 and 12 for specimens 4130N_05_2 and VAX300_06_2. During these two tests, visual observations were made, and it was noticed that the strain gauges were detached after the first few load cycles because of the small size of the glued area. Furthermore, the cyclic loading has also an effect on how quickly the gauges get detached and collect meaningless data.

It can be seen in Figs. 12 and 13 that the strain gauges attached to specimen 4130N_05_2 start collecting useless data after the first loading cycle, while the strain gauges of specimen VAX300_06_2 are detached approximately at the end of the fourth cycle. The strain data before the strain gauges were detached were used to calculate the actual applied stresses. It should be noted that the strains after that were considered to be constant and equal to the last meaningful values. The stress-strain relationships for the two specimens tested under cyclic tensile forces are shown in Figs. 14 and 15. It should be noted that the small spikes in Fig. 15 are caused when the applied loads (and selected target strain levels for the cycles) were being adjusted by the MTS testing machine.

Result of the Third Set of Experiments

Similar to the second phase, transverse strain values were also recorded in addition to the longitudinal strains in the third set of experiments. However, different

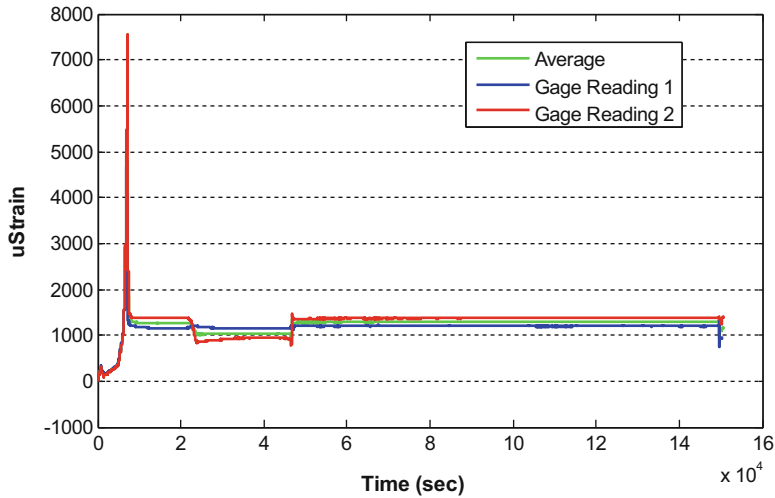


Fig. 12 Transverse strain-gauge readings of specimen 4130N_05_2

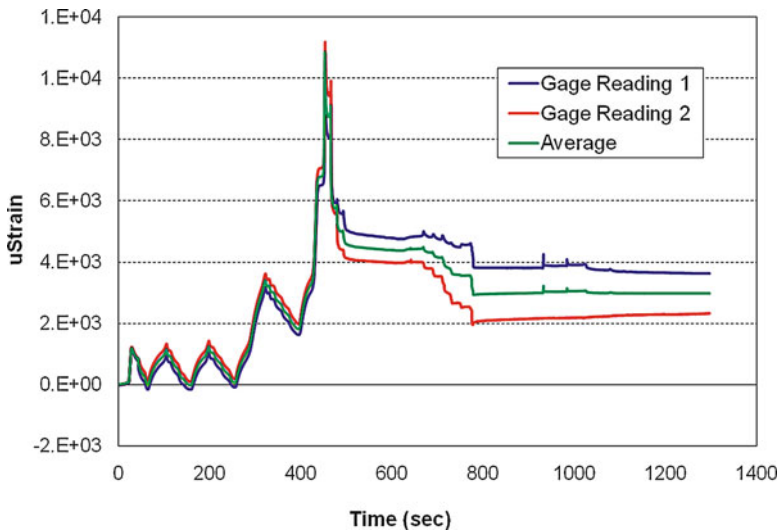


Fig. 13 Transverse strain-gauge readings of specimen VAX300_06_2

than the second set, a single strain gauge was attached to the midsection of one specimen from each material type (Fig. 16). The transverse strain along the thickness of the specimen was assumed to be equal to the transverse strain along the width.

The specimens (1080_04_2, 4130N_05_3 and VAX300_06_3) were, again, loaded up to certain stress/strain levels, and then they were unloaded until the

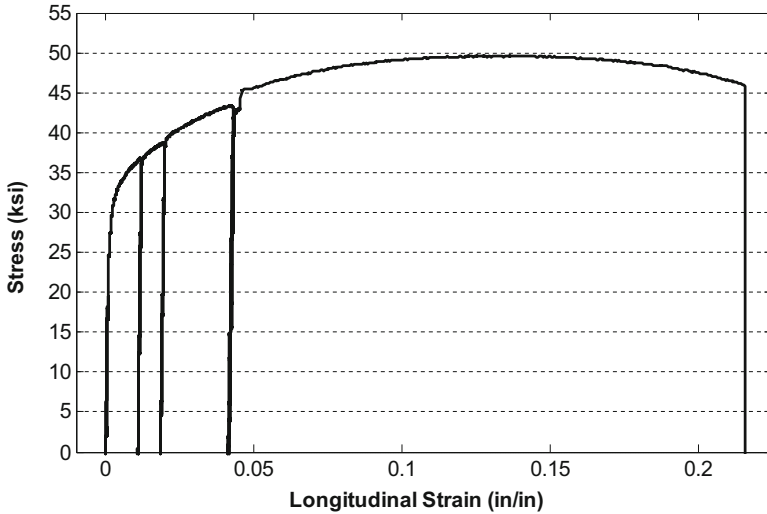


Fig. 14 Calculated stress-strain curve of specimen 4130N_05_2

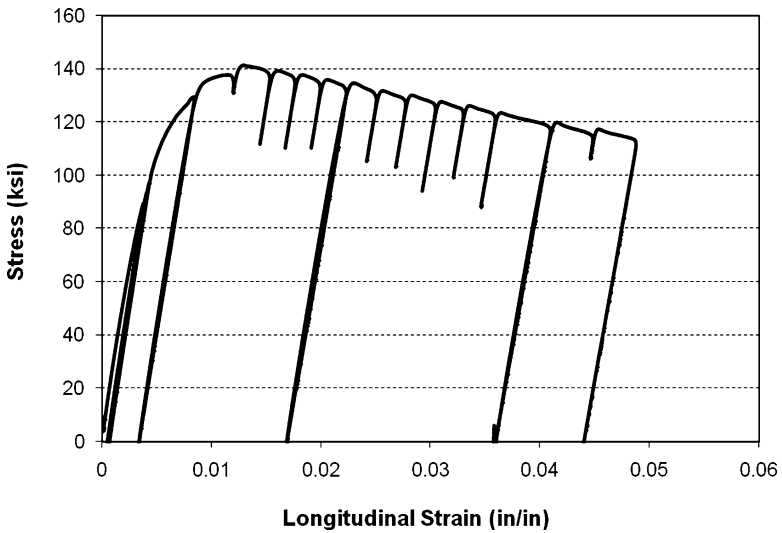


Fig. 15 Calculated stress-strain curve of specimen VAX300_06_2

applied force is equal to zero before the following loading. In order to calculate the applied stresses, the transverse strain data obtained from the single strain gauge were used. The strain recordings for all three tested specimens are presented in Figs. 17, 18, and 19. It can be seen from Fig. 17 that the strain gauge on the specimen 1080_04_2 kept recording valuable data until the end of the test while the ones

Fig. 16 A single strain gauge attached to the midsection of a test specimen

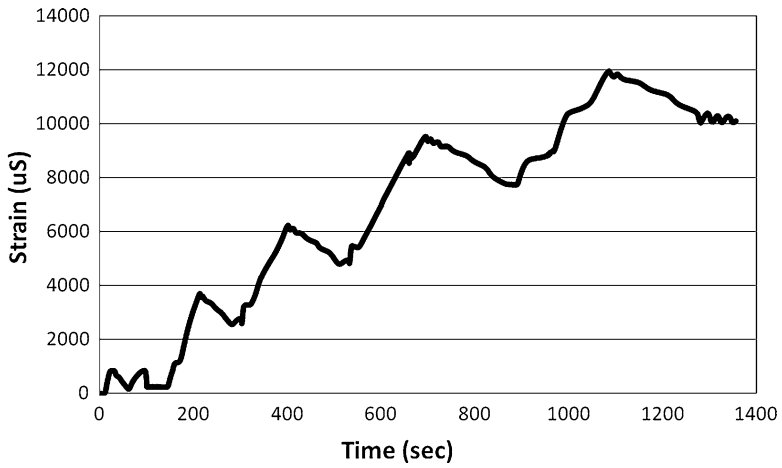
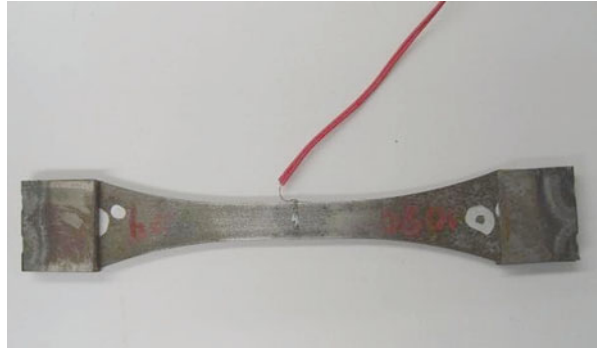


Fig. 17 Transverse strain-gauge readings of specimen 1080_04_2

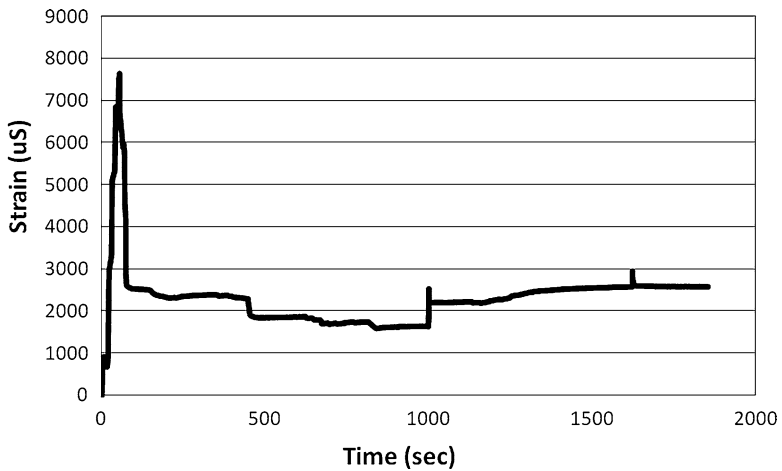


Fig. 18 Transverse strain-gauge readings of specimen 4130N_05_3

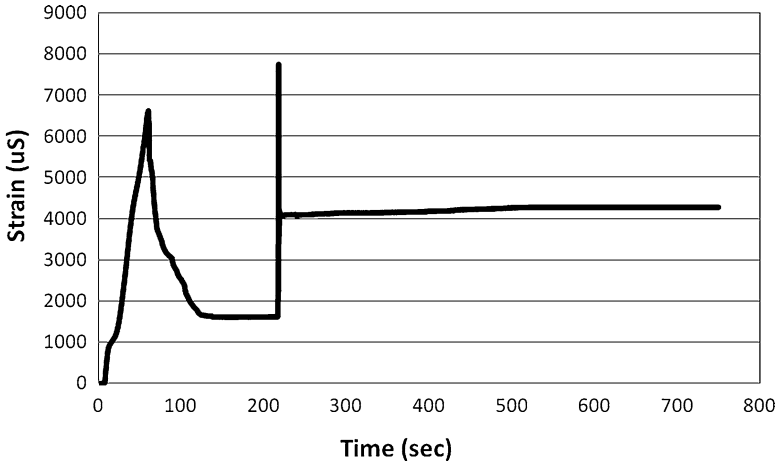


Fig. 19 Transverse strain-gauge readings of specimen VAX300_06_3

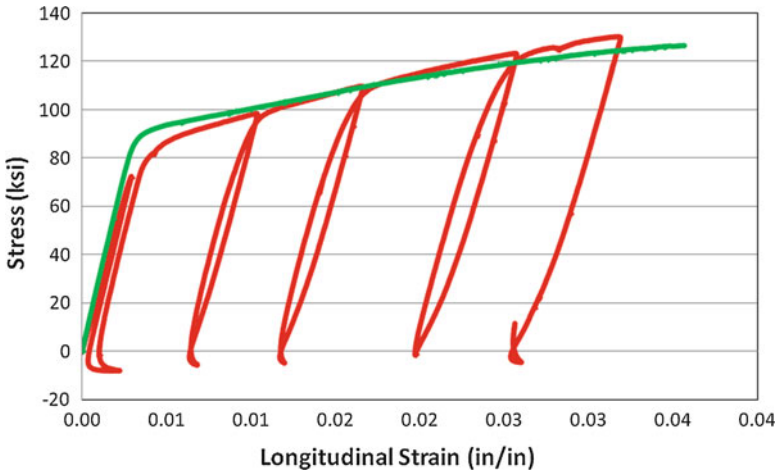


Fig. 20 Calculated stress-strain curve of specimen 1080_04_02 (red line) with the monotonic stress-strain curve of specimen 1080_04_1 (green line)

on 4130N and VascoMax C300 specimens detached by the end of the first loading cycle.

The meaningful strain data were used to calculate the actual applied stresses, and the stress-strain relationships for the three specimens tested under cyclic tensile forces are shown in Figs. 21 and 22 with red lines. The stress-strain relationships for the monotonic loading are also included in these figures with green lines (Fig. 20).

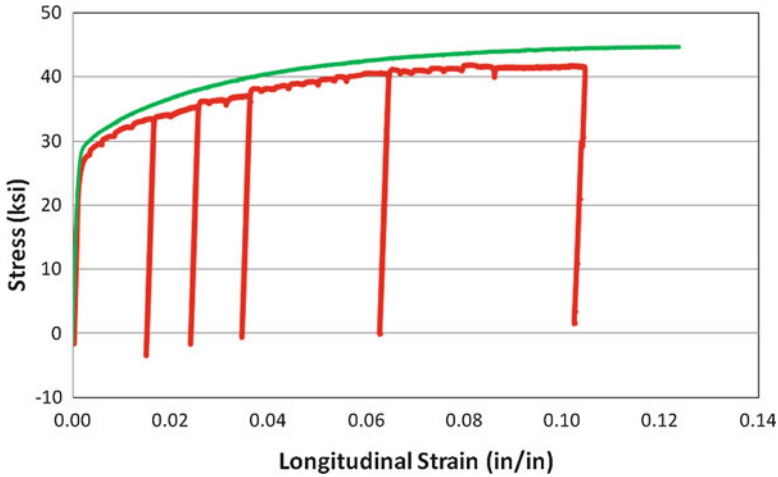


Fig. 21 Calculated stress-strain curve of specimen 4130N_05_03 (red line) with the monotonic stress-strain curve of specimen 4130N_05_1 (green line)

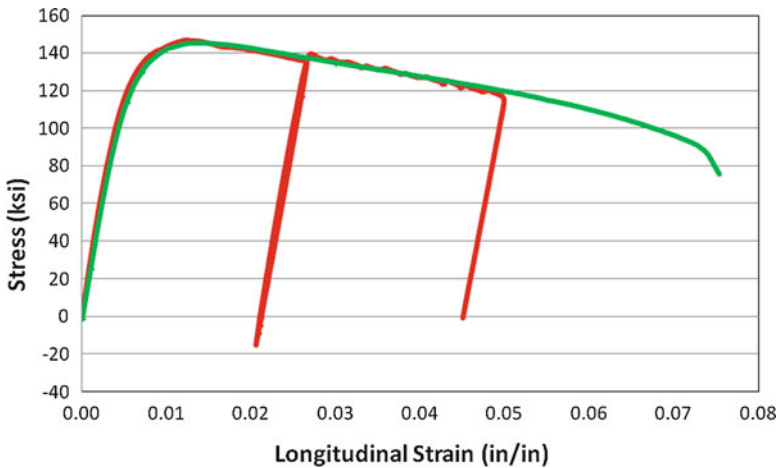


Fig. 22 Calculated stress-strain curve of specimen VAX300_06_03 (red line) with the monotonic stress-strain curve of specimen VAX300_06_1 (green line)

Numerical Simulations and Validating the Experimental Approach

The experimental work on the study of friction between metallic surfaces will be conducted by Voyiadjis and coworkers (Voyiadjis and Abu Al-Rub 2007; Voyiadjis et al. 2010; Voyiadjis and Deliktas 2009a, b; Lodygowski et al. 2011) with its counterpart at Ecole de' Nationale Institut der Mechanic (ENIM), at Metz, France. A modified Hopkinson's bar experiment shown schematically in Fig. 23a, developed

by Phillipon at ENIM, simulates high-velocity contact between two surfaces. In this experiment, ultrahigh-speed projectiles of the 1020 steel (for sled shoe surface) will be fired onto a stationary VASOMAX steel surface (for the sled track). The contact force in this experiment is measured by a dynamometer ring shown in Fig. 23b. The contact force exerted on the dynamometer (in Fig. 23c) can be predicted by measuring the change in length of the friction device using strain. The results from the experiment will include (a) the measured contact force (both normal and frictional) response from the dynamometer ring and (b) the final contact surface between the two specimens. The latter (along with the initial surface) will be

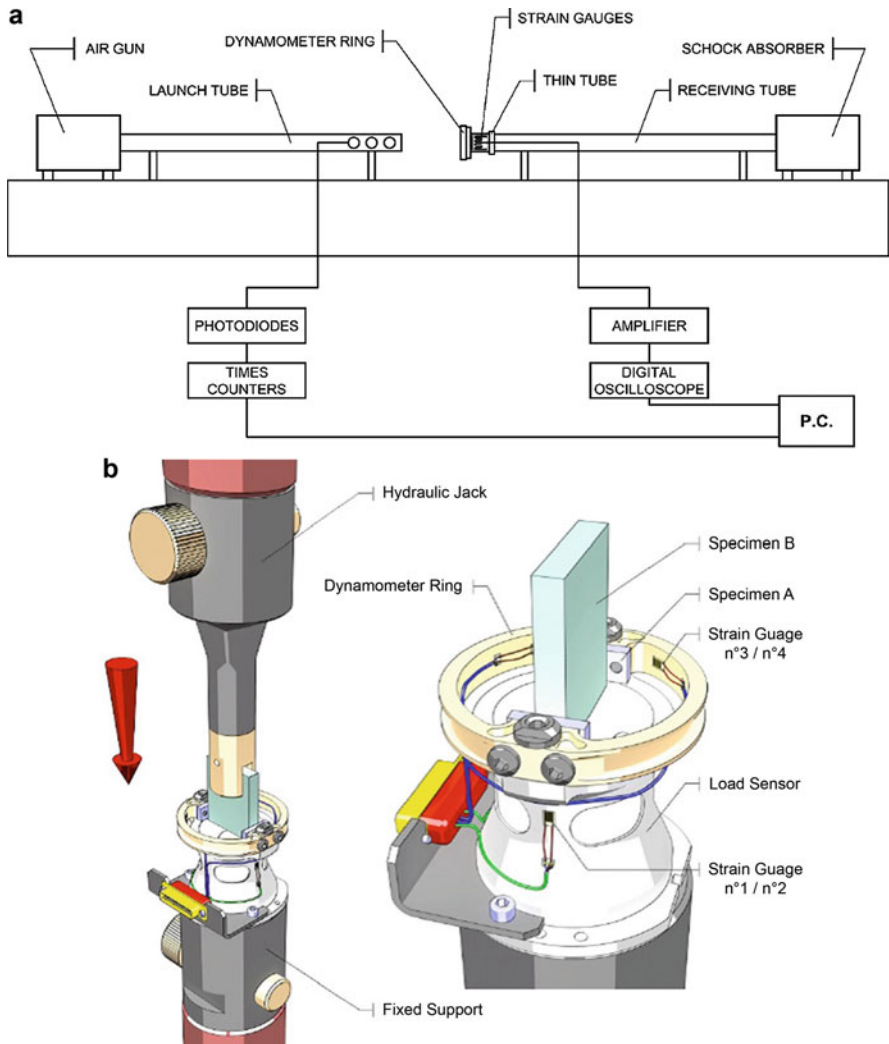


Fig. 23 (a) Schematic of modified Hopkinson experiment, developed by Phillipon et al. at (b) zoom-in view of friction device used to experimentally measure contact forces during high-velocity contact and (c) measured contact force on dynamometer

processed using SEM to characterize the topography of the surfaces before/after and consequently the size of contacting asperities. In this experiment, roughness measured through the SEM is characterized by software package (Vision32) provided by Veeco Metrology Group. The presliding characterization of the sled shoe and track surfaces indicates a presliding surface roughness of 4–8 mm, and it is postulated that damage phenomena are generated from dislocation generation at between crystallographically misoriented single crystals (at the microscale). In order to further characterize the sliding interfaces, experiments were conducted.

ENIM, used to measure contact forces during high-velocity contact between polycrystalline surfaces.

The simulations were performed with the FE commercial software ABAQUS (REFERENCE) according to the theory developed and coded as a VUMAT subroutine. The set of simulation was performed for different values of normal force and sliding velocity capturing the development of surface temperature as well as plastic deformation. The different mesh sizes were also considered to present the mesh independency thanks to the implemented theory. In the following figures, a couple of main results are presented.

In Fig. 24 the following stages of the simulation are presented using the coarse and fine mesh approach. The initial stage (Figs. 24a and 25a) presents the

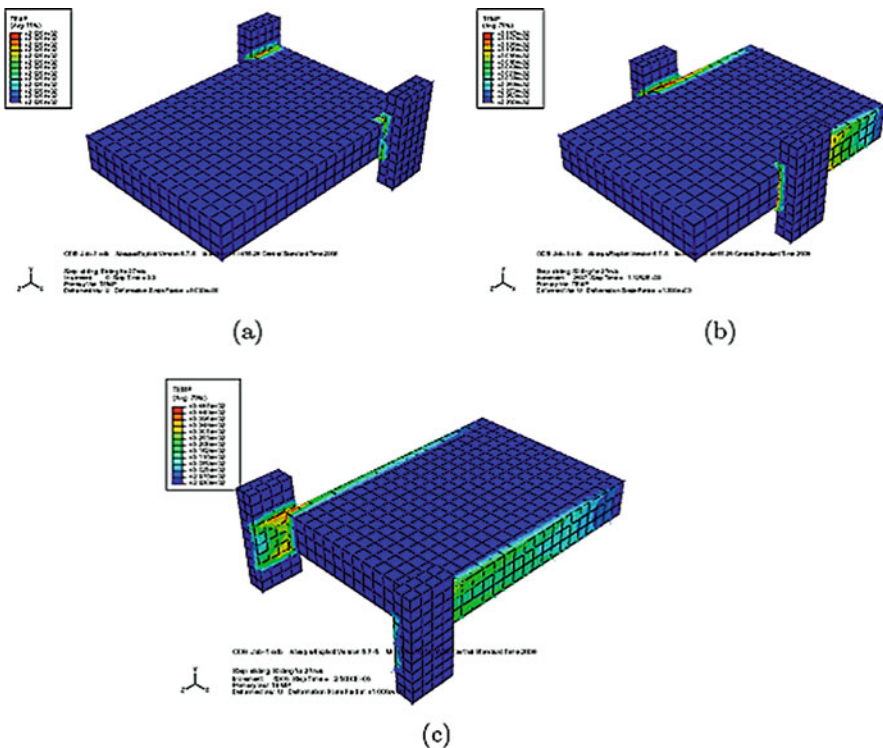


Fig. 24 FEM with coarse mesh and fine mesh for high-speed friction experiment with (a) boundary condition, (b) development temperature, (c) development of the moment

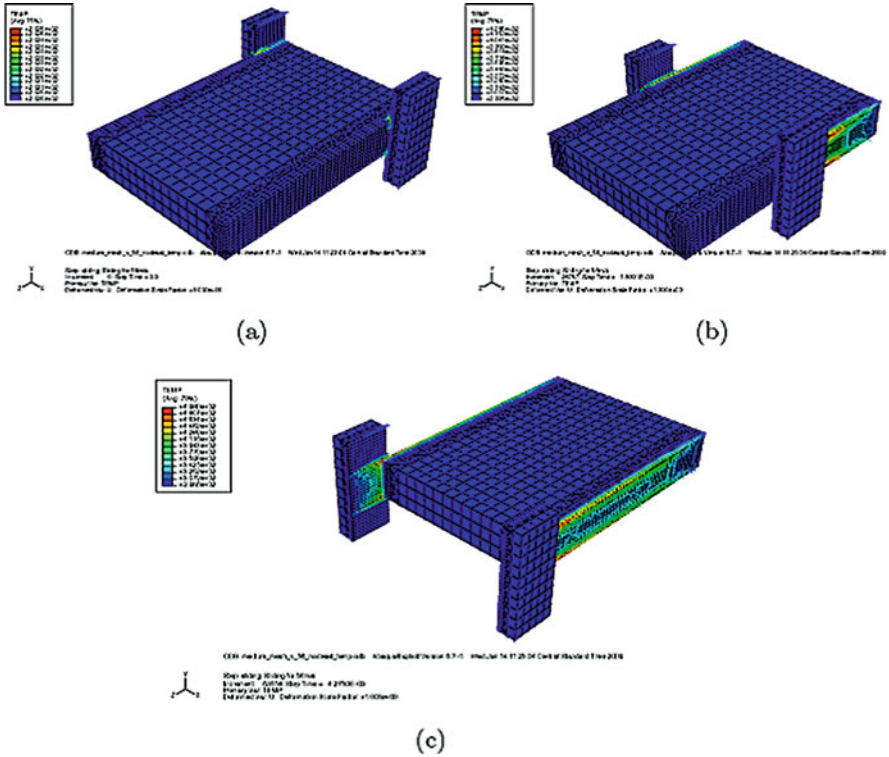


Fig. 25 FEM with fine mesh for high-speed friction experiment with (a) boundary condition, (b) development temperature, (c) development of the moment

application of all boundary conditions just before applying the sliding velocity to the specimen B. Figures 24b and 25b present the development of the temperature as one of the mainly considered variables. Lastly Figs. 24c and 25c show the moment when the contact between the specimens is lost. Figure 24 with the coarse mesh presents the results for the sliding velocity of 27 m=s, whereas Fig. 25 presents the fine mesh for the sliding velocity 56 m=s. Figure 24a presents the initial configuration of the simulation just before the sliding begins. Figure 24b presents the middle stage of the simulation process and the surface temperature development. Figure 24c shows the final stage of the simulation.

Discussions and Conclusions

Metal-to-metal friction problem is addressed in this work and revisited, specifically in investigating the coefficient of the dry friction for steel in the high-velocity range. The following contributions are made by this study:

- A physical model based on the nonlocal crystal plasticity model is proposed to investigate the coefficient of the dry friction. Results are presented from the simulations of the *simple shear of a constrained crystalline strip with double slips* and *high-speed friction experiment*.
- It is shown that the material responses obtained from the simulation of the physically based constitutive model agree with the real behavior of the metals.
- This work investigates also the effects of two thermodynamic processes, energetic and dissipative, involved in strain gradient crystal plasticity. The improved theory by Voyiadjis and Deliktas (2009b) is enhanced by decomposition for each thermodynamic force. This decomposition, in turn, introduces two different types of material length scales which provide better description of the dislocation network. The theoretical model is formulated for the crystalline system with multiple symmetric double slip systems. The boundary value problem analyzed here is the case of simple shear of a constrained crystalline strip in order to investigate the effects of the dissipative and energetic parameters on the size effect response of the constrained crystalline strip systems. The developed theory describes quantitatively the thickness of the boundary layer, hardening, and strengthening response of the system.
- However, there are still open issues of great interest in the field of SGP such as the following: (i) *dissipative and energetic nature of the dislocation network* (Anand et al. 2005; Bardella 2007; Bittencourt et al. 2003; Gurtin 2008; Voyiadjis and Deliktas 2009a) (the important question is how energy depends on the characteristics of the dislocation distribution), (ii) *evolution of the microstructure* which is attributed to many plasticity phenomena (Berdichevsky 2006; Gurtin 2008; Molinari and Ravichandran 2005; Nijs et al. 2008), and (iii) *physical bases of the material length scale* (Abu Al-Rub and Voyiadjis 2004; Anand et al. 2005; Mughrabi 2004).

References

- R.K. Abu Al-Rub, G.Z. Voyiadjis, Analytical and experimental determination of the material intrinsic length scale of strain gradient plasticity theory from micro- and nano-indentation experiments. *Int. J. Plast.* **20**, 1139–1182 (2004)
- K.E. Aifantis, J.R. Willis, Scale effects induced by strain gradient plasticity and interfacial resistance in periodic and randomly heterogeneous media. *Mech. Mater.* **38**, 702–716 (2006)
- L. Anand, M.E. Gurtin, S.P. Lele, C. Gething, A one-dimensional theory of strain-gradient plasticity: formulation, analysis, numerical results. *J. Mech. Phys. Solids* **53**, 1789–1826 (2005)
- K. Arakawa, Effect of time derivative of contact area on dynamic friction. *Appl. Phys. Lett.* **104**, 241603 (2014)
- K. Arakawa, An analytical model of dynamic sliding friction during impact. *Sci. Rep.* **7**, 40102 (2017)
- R.J. Asaro, Crystal plasticity. *J. Appl. Mech. Trans. Asme.* **50**, 921–934 (1983)
- L. Bardella, A deformation theory of strain gradient crystal plasticity that accounts for geometrically necessary dislocations. *J. Mech. Phys. Solids* **54**, 128–160 (2006)

- L. Bardella, Some remarks on the strain gradient crystal plasticity, with particular reference to the material length scales involved. *Int. J. Plast.* **23**, 296–322 (2007)
- E. Bayart, I. Svetlizky, J. Fineberg, Fracture mechanics determine the lengths of interface ruptures that mediate frictional motion. *Nat. Phys.* **12**, 166–170 (2016)
- V.L. Berdichevsky, On thermodynamics of crystal plasticity. *Scr. Mater.* **54**, 711–716 (2006)
- E. Bittencourt, A. Needleman, M.E. Gurtin, E. Van der Giessen, A comparison of nonlocal continuum and discrete dislocation plasticity predictions. *J. Mech. Phys. Solids* **51**, 281–310 (2003)
- R.J. Clifton, J. Duffy, K.A. Hartley, T.G. Shawki, On critical conditions for shear band formation at high strain rates. *Scr. Etall.* **18**, 443–448 (1984)
- N.A. Fleck, J.R. Willis, A mathematical basis for strain gradient plasticity theory. Part I: scalar plastic multiplier. *J. Mech. Phys. Solids* (2008). <https://doi.org/10.1016/j.jmps.2008.09.010>
- P. Fredriksson, P. Gudmundson, Competition between interface and bulk dominated plastic deformation in strain gradient plasticity. *Model. Simul. Mat. Sci. Eng.* **15**, S61–S69 (2007)
- M.E. Gurtin, On a framework for small-deformation viscoplasticity: free energy, microforces, strain gradient. *Int. J. Plast.* **19**, 47–90 (2003)
- M.E. Gurtin, A theory of grain boundaries that accounts automatically for grain misorientation and grain-boundary orientation. *J. Mech. Phys. Solids* **56**, 640–662 (2008)
- S. Hernandez, J. Hardell, H. Winkelmann, M.R. Ripoll, B. Prakash, Influence of temperature on abrasive wear of boron steel and hot forming tool steels. *Wear* **338–339**, 27–35 (2015)
- P. Ireman, Q.S. Nguyen, Using the gradients of temperature and internal parameters in continuum thermodynamics. *C. R. Mecanique* **332**, 249–255 (2004)
- P. Ireman, A. Klarbring, N. Stromberg, A model of damage coupled to wear. *Int. J. Solids Struct.* **40**, 2957–2974 (2003)
- L. Johansson, A. Klarbring, Thermoelastic frictional contact problems: modelling, FE-approximation and numerical realization. *Comput. Methods Appl. Mech. Eng.* **105**, 181–210 (1993)
- L. Johansson, A. Klarbring, Study of frictional impact using a nonsmooth equations solver. *J. Appl. Mech. Trans. ASME* **67**, 267–273 (2000)
- R.L. Johnson, M.A. Swikert, E.E. Bisson, Friction at high sliding velocities; naca-tn-1442 (1947)
- A. Klarbring, A mathematical-programming approach to 3-dimensional contact problems with friction. *Comput. Methods Appl. Mech. Eng.* **58**, 175–200 (1986)
- A. Klarbring, Examples of nonuniqueness and nonexistence of solutions to quasi-static contact problems with friction. *Ingenieur Arch.* **60**, 529–541 (1990)
- A. Lodygowski, G.Z. Voyiadjis, B. Deliktas, A. Palazotto, Non-local and numerical formulations for dry sliding friction and wear at high velocities. *Int. J. Plast.* **27**, 1004–1024 (2011)
- A. Molinari, G. Ravichandran, Constitutive modeling of high-strain-rate deformation in metals based on the evolution of an effective microstructural length. *Mech. Mater.* **37**, 737–752 (2005)
- H. Mughrabi, On the current understanding of strain gradient plasticity. *Mat. Sci. Eng. A Struct. Mat. Prop. Microstruct. Process.* **387–89**, 209–213 (2004)
- O. Nijs, B. Holmedal, J. Friis, E. Nes, Sub-structure strengthening and work hardening of an ultra-fine grained aluminum-magnesium alloy. *Mat. Sci. Eng. A Struct. Mat. Prop. Microstruct. Process.* **483**, 51–53 (2008)
- J.F. Nye, Some geometrical relations in dislocated crystals. *Acta Metall.* **1**, 153–162 (1953)
- M. Varga, M. Rojacz, H. Winkelmann, H. Mayer, E. Badisch, Wear reducing effects and temperature dependence of tribolayers formation in harsh environment. *Tribol. Int.* **65**, 190–199 (2013)
- G.Z. Voyiadjis, R.K. Abu Al-Rub, Nonlocal gradient-dependent thermodynamics for modeling scale-dependent plasticity. *Int. J. Multiscale Comput. Eng.* **5**, 295–323 (2007)
- G.Z. Voyiadjis, B. Deliktas, Formulation of strain gradient plasticity with interface energy in a consistent thermodynamic framework. *Int. J. Plast.* **25**(10), 1997–2024 (2009a)

-
- G.Z. Voyiadjis, B. Deliktas, Mechanics of strain gradient plasticity with particular reference to decomposition of the state variables into energetic and dissipative components. *Int. J. Eng. Sci.* **47**(11–12), 1405–1423 (2009b)
- G.Z. Voyiadjis, B. Deliktas, D. Faghihi, A. Lodygowski, Friction coefficient evaluation using physically based viscoplasticity model at the contact region during high velocity sliding. *Acta Mech.* **213**, 39–52 (2010)



Strain Gradient Plasticity: Deformation Patterning, Localization, and Fracture

27

Giovanni Lancioni and Tuncay Yalçinkaya

Contents

Introduction	972
Nonlocal Rate-Dependent and Rate-Independent Models	975
Problem Statement	975
Model Equations: Dissipation, Equilibrium, and Evolution	976
Rate-Dependent Model	979
Rate-Independent Model	981
Analytical Solutions	983
RD Model	983
RI Model	984
Numerical Results	988
Plastic Slip Patterning	988
Necking in Tensile Steel Bars	992
Conclusions	997
References	998

Abstract

In this chapter, two different strain gradient plasticity models based on non-convex plastic energies are presented and compared through analytical estimates

G. Lancioni (✉)

Dipartimento di Ingegneria Civile, Edile e Architettura, Università Politecnica delle Marche,
Ancona, Italy

e-mail: g.lancioni@univpm.it

T. Yalçinkaya

Aerospace Engineering Program, Middle East Technical University Northern Cyprus Campus,
Guzelyurt, Mersin, Turkey

Department of Aerospace Engineering, Middle East Technical University, Ankara, Turkey

e-mail: yalcinka@metu.edu.tr

and numerical experiments. The models are formulated in the simple one-dimensional setting, and their ability to reproduce heterogeneous plastic strain processes is analyzed, focusing on strain localization phenomena observed in metallic materials at different length scales. In a geometrically linear context, both models are based on the additive decomposition of the strain into elastic and plastic parts. Moreover, they share the same non-convex plastic energy, and they are both characterized by the same nonlocal plastic energy as well, i.e., a quadratic form of the plastic strain gradient. In the first model, proposed in Yalçinkaya et al. (*Int J Solids Struct* 49:2625–2636, 2012) and Yalcinkaya (*Microstructure evolution in crystal plasticity: strain path effects and dislocation slip patterning*. Ph.D. thesis, Eindhoven University of Technology, 2011), the plastic energy is assumed to be conservative, and plastic dissipation is introduced through a viscous term, which makes the formulation rate-dependent. In the second model, developed in Del Piero et al. (*J Mech Mater Struct* 8(2–4):109–151, 2013), the plastic term is supposed to be totally dissipative. As a result, plastic deformations are not recoverable, and the resulting framework is rate-independent, contrary to the first model. First, the evolution problems resulting from the two theories are analytically solved in a special simplified case, and correlations between the shape of the plastic potential and the modeling predictions are established. Then, the models are numerically implemented by finite elements, and numerical solutions of two different one-dimensional problems, associated with different plastic energies, are determined. In the first problem, a double-well plastic energy is considered, and the evolution of plastic slip patterning observed in crystals at the mesoscale is reproduced. In the second problem, a convex-concave plastic energy is used to simulate the macroscopic response of a tensile steel bar, which experiences the so-called necking process, with plastic strains localization and final coalescing into fracture. Numerical results provided by the two models are analyzed and compared.

Keywords

Strain gradient plasticity · Size effect · Localization · Deformation patterning · Damage · Fracture

Introduction

In recent years, developments in constitutive modeling have illustrated the ability of strain gradient plasticity theories to capture different strain localization processes by incorporating non-convex plastic energies into the model in a thermodynamically consistent way (see, e.g., Yalçinkaya et al. 2011, 2012; Klusemann and Yalçinkaya 2013; Klusemann et al. 2013 for the modeling of intragranular microstructure evolution resulting in plastic anisotropy and Del Piero et al. 2013; Lancioni 2015 for strain localization leading to failure). Since processes like strain localization and microstructures evolution are stress-softening processes, the corresponding evolution boundary value problems undergo convergence troubles due to loss

of ellipticity. In order to remedy these problems, several methods have been proposed including variational regularization methods, nonlocal methods, viscous regularization techniques, and Cosserat theories. However, nowadays, the modeling of softening mechanisms is still an open issue, and it is subject to intense research activities. In order to contribute to this, in this chapter two strain gradient plasticity models incorporating the phase-field idea for the evolution of plastic slip are proposed and analyzed. The numerical comparison study of rate-dependent and rate-independent models has been addressed recently in Yalçinkaya and Lancioni (2014) and Lancioni et al. (2015a). In the current chapter, we focus on the similarities and differences of both models in a variational format and present analytical and numerical solutions for that purpose. Models are formulated in the simple one-dimensional setting, which preserves simplicity and allows to focus on the key physical and mathematical aspects, avoiding the complicated technicalities of multidimensional frameworks, and the hypothesis of linear elastoplasticity is assumed, which allows to decompose the total strain into the sum of elastic and plastic contributions. Extensions to multidimensional setting, which would result in (strain gradient) crystal plasticity models (see, e.g., Yalçinkaya et al. 2012; Lancioni et al. 2015b) and/or finite elasticity frameworks, are not included.

The first model, proposed in Yalçinkaya et al. (2011, 2012), is deduced by following a classical thermodynamical approach (see Gurtin et al. 2010). The principle of virtual work is used to get balance equations at macro and micro level, while the dissipation inequality is exploited to define micro-stresses. In order to satisfy the dissipation inequality, a viscous plastic stress is introduced, which makes the model rate-dependent (RD). The free energy is assumed to be sum of a non-convex plastic term (which does not exist in Gurtin et al. 2010) and two quadratic terms with respect to the elastic deformation and the plastic deformation gradient. The inclusion of a non-convex plastic term in free energy that drives the localization and patterning is arguable, yet it is physical for the illustration of deformation patterning under monotonic loading, which is the main scope of the current work.

In the second model, developed in Del Piero et al. (2013), discussed in Del Piero (2013), and extended in Lancioni (2015), the plastic evolution is determined by incremental minimization of an energy functional, which is composed by three contributions equal to those considered in the abovementioned free energy of the RD model. However, in this case, the plastic term is supposed to be totally dissipative, and, therefore, plastic deformations are not recoverable. The resulting framework is rate-independent (RI), contrary to the previous model.

The different dissipative properties assigned to the plastic energy in two formulations reflect on different descriptions of the plastic strain evolution. In the RD model, plastic strains are partially recoverable, because of the plastic free energy term, and partially dissipate, since the viscous stress contributes to dissipate plastic power. On the contrary, in the RI model, plastic strains are totally irreversible, being the corresponding energy dissipative. In both the two models, stability issues related to stress-softening regimes are solved by incorporation of the plastic strain gradient term into the expression of the free energy. This nonlocal term not only allows to regularize the evolution problem but also plays the fundamental role of

strain localization limiter against abrupt fractures (see Bazant and Jirásek 2002). Moreover, it also gives the opportunity to work at different length scales due to the internal length scale parameter included intrinsically in its formulations.

The aim of this study is to analyze the influence of the form of the plastic energy function on the description of heterogeneous plastic processes. In this respect, analytical solutions are found in the simplified case of evolution from homogeneous strain configurations, providing precise indications on the shape to give to the plastic energy in order to reproduce specific evolution processes. Then, finite element codes are developed by implementing the RD and RI models and used to solve numerically two different problems of strain localization occurring at different length scales, as described in the following:

- (i) The first problem is the evolution of heterogeneous plastic shear strains in metallic materials. When metals are subject to considerable loadings (e.g., during forming processes), mechanisms of dislocation self-organization activate that lead to the formation of patterning observable at the mesoscale, where regions of high dislocation density (dislocation walls) envelop areas of low dislocation density (dislocation cell interiors), which can also be regarded as domains of high plastic slip and low plastic slip activity, as shown in the picture of Fig. 1a. To reproduce this patterning process, the plastic energy is assumed to be a Landau-Devonshire type of double-well functional, where the second well is shifted up, similar to the form used in Yalçinkaya et al. (2011). Using such a type of plastic potential results in a Ginzburg-Landau phase-field-like relation for the evolution of plastic slip, where the different phases are identified as regions with high plastic and low plastic strain.

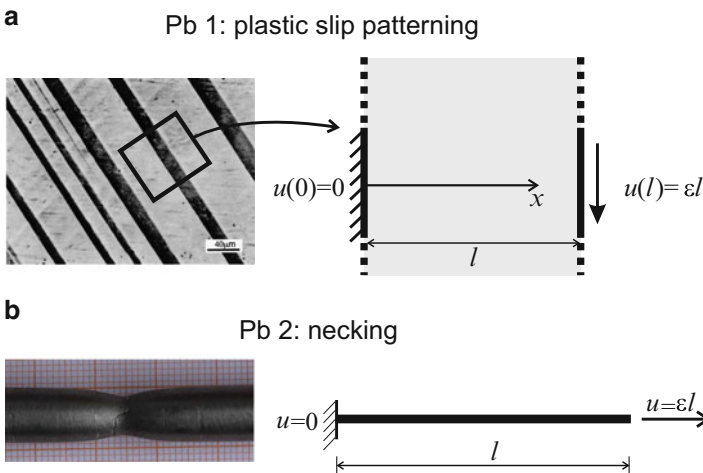


Fig. 1 Plastic slip patterning: (a) straight slip bands on a single crystal of 3.25% silicon iron (Hull 1963); (b) one-dimensional geometrical scheme of the shear problem. Necking: (c) neck in a tensile steel bar; (d) one-dimensional scheme of a tensile bar

- (ii) The second problem refers to the process of strain localization that leads to final fracture, observed in tensile metals, the so-called necking process (see Fig. 1c). Necking occurs at a length scale which is almost at the macroscopic engineering level, thus completely different from that of the previous problem. For this problem, a convex-concave plastic functional has been chosen. Localization initiates and evolves when the amount of plastic deformation reaches values in the concave part of the energy, and it is accompanied by substantial stress-softening. Localization continues on smaller and smaller regions, until final fracture, which corresponds to the solution loss of stability. The response of steel bars to tensile loadings has been studied in Lancioni (2015), by using an enriched version of the RI model. In the present study, it is solved by the RD model as well, and comparisons are established.

The chapter is organized as follows. In section “Nonlocal Rate-Dependent and Rate-Independent Models,” the two models are formulated within a common unified framework, which allows to easily point out analogies and differences. In section “Analytical Solutions,” analytical solution is found, and in section “Numerical Results” numerical results are presented and commented. Concluding remarks are summarized in section “Conclusions.”

Nonlocal Rate-Dependent and Rate-Independent Models

Problem Statement

Consider a one-dimensional domain $(0, l)$ of length l . The displacement of a point $x \in (0, l)$ at the time instant t is denoted by $u = u_t(x)$. Here and in the following, the dependence on time is indicated by a subscript, and the following notation for derivatives is used: for any function $v = v_t(w)$, which depends on a certain space-dependent variable $w = w(x)$ and on time t , a prime indicates derivative with respect to w , $v' = dv/dw$, and a dot means time derivative, $\dot{v} = dv/dt$.

The boundary conditions assigned at the domain endpoints are

$$u_t(0) = 0, \quad u_t(l) = l\varepsilon_t, \quad (1)$$

where ε_t is an imposed deformation, function of time. We suppose that the deformation is decomposed additively into an elastic part ε^e and a plastic part ε^p , i.e.,

$$u' = \varepsilon^e + \varepsilon^p. \quad (2)$$

Since positive deformations are applied ($\varepsilon_t \geq 0$), the plastic deformation coincides with the cumulative plastic strain, here denoted by γ . Thus we can write

$$u' = \varepsilon^e + \gamma, \quad (3)$$

In addition to (1), two possible boundary conditions on γ can be assigned, which are

$$\gamma_i(0) = \gamma_i(L) = 0 \text{ (hard b.c.)}, \text{ or } \gamma'_i(0) = \gamma'_i(L) = 0 \text{ (soft b.c.)}. \quad (4)$$

As shown in the numerical simulations, hard boundary conditions force the plastic strain γ to evolve in the central part of the body, and a boundary layer develops at the both ends of the bar. On the other hand, soft boundary conditions lead to evolution of a homogeneous γ field. An interesting discussion is proposed in Jirásek and Rolshoven (2009) on possible assignment of boundary conditions for the plastic strain in gradient plasticity models.

Model Equations: Dissipation, Equilibrium, and Evolution

Both theories are formulated within a common unified framework, which allows an easy comparison from a theoretical viewpoint. In here, the general modeling ingredients are introduced that are needed to deduce balance and evolution equations in a thermodynamically consistent way. In the next sections “[Rate-Dependent Model](#)” and “[Rate-Independent Model](#),” these relations are specialized for the models, according to the specific constitutive assumptions of each theory.

Given the triplet (u, γ, γ') of independent internal variable, the material is assumed to be endowed with the free energy

$$\psi(\varepsilon^e, \gamma, \gamma') = \psi_e(\varepsilon^e) + \psi_\gamma(\gamma) + \psi_{\gamma'}(\gamma'), \quad (5)$$

which is sum of elastic, plastic, and stored contributions, analogously to the free energy considered in Gurtin and Anand (2009). Here, we assume that the elastic and the nonlocal terms are quadratic functions

$$\psi_e(\varepsilon^e) = \frac{1}{2} E \varepsilon^{e2}, \quad \psi_{\gamma'}(\gamma') = \frac{1}{2} A \gamma'^2, \quad (6)$$

but, in general, different expressions can be considered. The plastic potential ψ_γ is a monotonic increasing function of γ , which can assume any expression. The macroscopic stress and the microscopic stress and hyperstress power-conjugated to $\dot{\varepsilon}^e$, $\dot{\gamma}$ and $\dot{\gamma}'$, respectively, are defined as

$$\sigma = \psi'_e(\varepsilon^e) = E \varepsilon^e, \quad \pi = \psi'_\gamma(\gamma), \quad \xi = \psi'_{\gamma'}(\gamma') = A \gamma'. \quad (7)$$

These stresses are energetic. The next stage concerns the definition of the local dissipated power.

Dissipation. If we suppose that dissipation is only due to plastic strains, the dissipated power has the form

$$D(\gamma, \dot{\gamma}) = \sigma^d(\gamma, \dot{\gamma})\dot{\gamma}, \quad (8)$$

where $\sigma^d = \sigma^d(\gamma, \dot{\gamma})$ is a dissipated micro-stress power-conjugated to $\dot{\gamma}$. The dependence of σ^d on the rate $\dot{\gamma}$ allows to introduce viscous plastic micro-stresses within the model, as done in the RD model of section “[Rate-Dependent Model](#).” Since the dissipation inequality for isothermal processes reads

$$D = P - \dot{\psi} \geq 0, \quad (9)$$

with P the local internal power, σ^d must be such that the dissipation D in (8) is non-negative. A further way to introduce dissipative behavior is through a dissipation potential, according to the theory of generalized standard materials (Mielke 2006). If $\theta(\gamma)$ is a dissipative plastic potential, the dissipated micro-stress is $\theta'(\gamma)$, and the dissipation inequality (9) becomes

$$D(\gamma, \dot{\gamma}) = \theta'(\gamma)\dot{\gamma} \geq 0. \quad (10)$$

Thus, summing up, the total micro-stress conjugated to γ is

$$\sigma_\gamma = \pi + \sigma^d + \theta', \quad (11)$$

where the first term is energetic, and the second and third terms are dissipated. If we assume that volume forces are null, the body total energy is

$$E(u, \gamma) = \int_0^l (\psi(u' - \gamma, \gamma, \gamma') + \theta(\gamma)) dx, \quad (12)$$

which is used in the following to deduce the equilibrium and the evolution equations in a variational format.

Equilibrium. A configuration (u, γ) is equilibrated if

$$\delta E(u, \gamma; \delta u, \delta \gamma) + W_{NC}(\gamma, \dot{\gamma}; \delta \gamma) \geq 0, \quad (13)$$

for any admissible perturbation $(\delta u, \delta \gamma)$, where δE is the first variation of E , and W_{NC} is the infinitesimal virtual work of σ^d . Using (7), the two terms in the above inequality are

$$\begin{aligned} \delta E &= \int_0^l (\sigma \delta u' + (\pi + \theta' - \sigma)\delta \gamma + \xi \delta \gamma') dx, \\ W_{NC} &= \int_0^l \sigma^d \delta \gamma dx. \end{aligned} \quad (14)$$

When the unknown pair (u, γ) is not constrained through inequality conditions, the equilibrium condition (13) reduces to an equality. Indeed, in the case of unconstrained (u, γ) , for any admissible perturbation $(\delta u, \delta \gamma)$, also the opposite perturbation $(-\delta u, -\delta \gamma)$ is admissible, and thus (13) must be satisfied as an equality, in order to be satisfied for any pair of opposite perturbations. The resulting equation is the equation of virtual power (see Gurtin et al. 2010). This situation is found in the RD model. Differently, in the RI model, an inequality constrain on γ leads to an equilibrium inequality.

Evolution. The evolution problem is solved by following a variational incremental procedure, which allows to determine the solution $(u_{t+\tau}, \gamma_{t+\tau})$ at the time instant $t + \tau$, with τ a given time increment, when the solution (u_t, γ_t) at the previous instant t is known.

The solution at $t + \tau$ is approximated by the first-order Taylor expansion

$$u_{t+\tau} = u_t + \tau \dot{u}_t, \quad \gamma_{t+\tau} = \gamma_t + \tau \dot{\gamma}_t, \tag{15}$$

where the velocity pair $(\dot{u}_t, \dot{\gamma}_t)$ represents the unknown to be determined. The total energy (12) is approximated by the second-order development

$$E_{t+\tau}(\dot{u}_t, \dot{\gamma}_t) = E_t + \tau \dot{E}_t(\dot{u}_t, \dot{\gamma}_t) + \frac{1}{2} \tau^2 \ddot{E}_t(\dot{u}_t, \dot{\gamma}_t), \tag{16}$$

and

$$\Delta E_t(\dot{u}_t, \dot{\gamma}_t) = E_{t+\tau} - E_t = \tau \dot{E}_t(\dot{u}_t, \dot{\gamma}_t) + \frac{1}{2} \tau^2 \ddot{E}_t(\dot{u}_t, \dot{\gamma}_t) \tag{17}$$

is the approximated energy increment within a time step $t \rightarrow t + \tau$. The pair $(\dot{u}_t, \dot{\gamma}_t)$ solves the variational inequality

$$\delta \Delta E_t(\dot{u}_t, \dot{\gamma}_t; \delta \dot{u}_t, \delta \dot{\gamma}_t) + \Delta W_{NC}(\dot{\gamma}_t; \delta \dot{\gamma}_t) \geq 0, \tag{18}$$

for any admissible perturbations pair $(\delta \dot{u}_t, \delta \dot{\gamma}_t)$, where $\Delta W_{NC}(\dot{\gamma}_t; \delta \dot{\gamma}_t)$ is the virtual power expended by σ^d within the time interval $(t, t + \tau)$, i.e.,

$$\Delta W_{NC} = \int_0^t (\sigma_{t+\tau}^d - \sigma_t^d) \delta \dot{\gamma} \, dx. \tag{19}$$

As above commented for the equilibrium relation (13) and (18) reduces to an equality, when the unknowns $(\dot{u}_t, \dot{\gamma}_t)$ are not constrained by inequality conditions.

The formulation proposed in this section allows to obtain equilibrium and evolution equations, provided that certain constitutive assumptions are made. In particular, the quantities to be assigned are the plastic free energy $\psi_\gamma(\gamma)$, the dissipative micro-stress σ^d , and the dissipation potential $\theta(\gamma)$, which must satisfy the dissipation inequality (10). We recall that the elastic and the nonlocal

free energies are assigned in (6). In the next sections “[Rate-Dependent Model](#)” and “[Rate-Independent Model](#),” specific constitutive assumptions are made, and both models are formulated.

Rate-Dependent Model

In this section, we formulate the RD model, which was first proposed in Yalçinkaya et al. (2011) by using the principle of virtual power.

Constitutive assumptions. The basic constitutive assumptions are:

1. The plastic free energy $\psi_\gamma(\gamma)$ satisfies the conditions $\psi_\gamma(0) = 0$, $\psi'_\gamma(0) \geq 0$, and $\psi'_\gamma(\gamma) > 0$ for any $\gamma > 0$. Two different non-convex expressions of $\psi_\gamma(\gamma)$ will be given for the two problems numerically solved in sections “[Plastic Slip Patterning](#)” and “[Necking in Tensile Steel Bars](#).”
2. A viscous dissipative micro-stress of the form

$$\sigma^d = c\dot{\gamma} \quad (20)$$

is assumed, where c is a viscous coefficient. Notice that, with this stress, the dissipation inequality (10) is automatically satisfied. The reader is referred to Yalçinkaya et al. (2011) for a discussion on more complex expressions of σ^d .

3. The dissipative potential $\theta(\gamma)$ is neglected.

With these assumptions, the micro-stress power-conjugated to $\dot{\gamma}$ is

$$\sigma_\gamma = \psi'_\gamma + c\dot{\gamma}, \quad (21)$$

where the first term is energetic, and the second is dissipative. The viscous contribution constitutes the source of rate dependency of the model.

Equilibrium. Under these assumptions, the equilibrium condition (13) rewrites

$$\int_0^l \left(\sigma(\delta u' - \delta\gamma) + \psi'_\gamma \delta\gamma + \xi \delta\gamma' + c\dot{\gamma} \delta\gamma \right) dx = 0, \quad (22)$$

for any admissible perturbation pair $(\delta u, \delta\gamma)$, from which the macroscopic and microscopic force balance equations are obtained by assuming $\delta\gamma = 0$ and $\delta u' = 0$ (i.e., $\delta\gamma = -\delta\varepsilon^e$), respectively. They are

$$\sigma' = 0, \quad \sigma - \psi'_\gamma + \xi' - c\dot{\gamma} = 0. \quad (23)$$

Remark 1. If we assume $\psi'_\gamma(0) > 0$, solution of the above equilibrium equations at the initial instant $t = 0$ is

$$\varepsilon_0^e = \frac{\psi'_\gamma(0)}{E}, \quad \gamma_0 = 0. \tag{24}$$

This initial strain corresponds to the prestress $\sigma_0 = \psi'_\gamma(0)$ associated with the pre-imposed deformation $\varepsilon_0 = \psi'_\gamma(0)/E$. From such an initial state, the elastoplastic deformation evolves according to the evolution equation

$$E\varepsilon^e - \left(\psi'_\gamma - \psi'_\gamma(0) \right) + A\gamma'' - c\dot{\gamma} = 0, \tag{25}$$

which differs from (23)₂ for the second term within the brackets, which is null at the beginning of the evolution process. The constitutive assumption $\psi'_\gamma(0) > 0$ is made to reproduce initial purely elastic behaviors.

Evolution. The evolution equation is obtained from (18), with

$$\Delta W_{NC} = \int_0^l (c\dot{\gamma} - \sigma_t^d)\delta\dot{\gamma} dx. \tag{26}$$

The macroscopic and microscopic evolution equations are determined by assuming $\delta\dot{\gamma} = 0$ and $\delta\dot{u}' = 0$, respectively, and they are

$$(\dot{u}'_t - \dot{\gamma}_t)' = 0, \quad E(\dot{u}'_t - \dot{\gamma}_t) - \left(\psi''_\gamma(\gamma_t) + \frac{c}{\tau} \right) \dot{\gamma}_t + A\dot{\gamma}''_t = -\frac{\sigma_t^d}{\tau}, \tag{27}$$

with the boundary conditions

$$\begin{aligned} \dot{u}_t(0) = 0, \quad \dot{u}_t(l) = l\dot{\varepsilon}_t, \\ \dot{\gamma}_t(0) = \dot{\gamma}_t(l) = 0 \text{ (hard b.c.)}, \quad \text{or} \quad \dot{\gamma}'_t(0) = \dot{\gamma}'_t(l) = 0 \text{ (soft b.c.)}. \end{aligned} \tag{28}$$

Equations (27) are obtained in Yalçinkaya et al. (2011) by linearizing equations (23). Equation (27)₁ states that $(\dot{u}'_t - \dot{\gamma}_t)$ is constant, and thus it can be written in the form

$$\dot{u}'_t - \dot{\gamma}_t = \frac{1}{l} \int_0^l (\dot{u}'_t - \dot{\gamma}_t) dx = \dot{\varepsilon}_t - \bar{\dot{\gamma}}_t, \tag{29}$$

where

$$\bar{\dot{\gamma}}_t = \frac{1}{l} \int_0^l \dot{\gamma}_t dx \tag{30}$$

is the mean value of $\dot{\gamma}_t$ in $(0, l)$. Substituting (29) in (27)₂, we get the following set of equations, alternative to (27),

$$\begin{aligned}
 A\dot{\gamma}_t'' - \left(\psi_\gamma''(\gamma_t) + \frac{c}{\tau} \right) \dot{\gamma}_t - E\bar{\gamma}_t &= -E\dot{\varepsilon}_t - \frac{\sigma_t^d}{\tau}, \\
 \dot{u}_t' &= \dot{\varepsilon}_t + \dot{\gamma}_t - \bar{\dot{\gamma}}_t,
 \end{aligned}
 \tag{31}$$

where the first equation depends only on $\dot{\gamma}_t$. Equations (31) are equivalent to (27), and they are solved by determining $\dot{\gamma}_t$ from the first equation and, then, \dot{u}_t from the second equation.

Rate-Independent Model

In this section, we deduce the governing equations of the RI model first proposed in Del Piero et al. (2013).

Constitutive assumptions. We make the following constitutive assumptions:

1. The plastic free energy $\psi_\gamma(\gamma)$ is neglected.
2. The dissipative micro-stress σ^d is neglected.
3. A dissipative potential $\theta(\gamma)$ is considered, which is assumed to be equal to the plastic free energy of the RD model, defined at point (1) of section “[Rate-Dependent Model](#).” Thus the dissipative potential must satisfy the conditions $\theta(0) = 0$, $\theta'(0) \geq 0$, and $\theta'(\gamma) > 0$, for any $\gamma > 0$.

With these hypotheses, the plastic micro-stress is

$$\sigma_\gamma = \theta'(\gamma), \tag{32}$$

which is totally dissipative, in agreement with Aifantis theory (Aifantis 1984) and the reformulations proposed in Gudmundson (2004) and Gurtin and Anand (2009) in a thermodynamically consistent format, where the local term of the flow rule is dissipative and the nonlocal term is energetic.

Since $\theta'(\gamma) \geq 0$ for any γ , according to the assumptions given in the above point 3, the dissipation inequality (10) is satisfied if

$$\dot{\gamma} \geq 0, \tag{33}$$

which represents a constraint for the plastic strain and imposes that plastic strain can never decrease in an evolution process.

Equilibrium. From (33) an admissible perturbation is such that $\delta\gamma \geq 0$. Since $W_{NC} = 0$, inequality (13) rewrites

$$\int_0^l (\sigma(\delta u' - \delta\gamma) + \xi\delta\gamma' + \theta'\delta\gamma) dx \geq 0, \tag{34}$$

for any admissible plastic strain perturbation such that $\delta\gamma \geq 0$. Assuming $\delta\gamma = 0$, we obtain the balance equation (23)₁, while, assuming $\delta u' = 0$, we found the plastic yield condition

$$\sigma \leq \theta' - \xi', \tag{35}$$

which states that the stress σ cannot be greater than the yield limit $\theta' - \xi'$.

Evolution. Evolution is governed by inequality (18), where $\Delta W_{NC} = 0$. Assuming $\delta\dot{\gamma} = 0$, we get the macroscopic evolution equation (27)₁, and, setting $\delta\dot{u}' = 0$, we obtain

$$\delta\Delta E_t(\dot{u}_t, \dot{\gamma}_t, 0, \delta\dot{\gamma}) = \int_0^l (f_t + \tau \dot{f}_t) \delta\dot{\gamma} \, dx + [(\xi_t + \tau \dot{\xi}_t) \delta\dot{\gamma}]_0^l \geq 0, \tag{36}$$

where

$$f_t = \theta'_t - \sigma_t - \xi'_t, \tag{37}$$

and $\delta\dot{\gamma}$ is an arbitrary perturbation that satisfies the condition

$$\dot{\gamma}_t + \delta\dot{\gamma} \geq 0, \tag{38}$$

imposed by the irreversibility relation (33). The boundary terms are null, if the boundary conditions (28) are assigned. For non-negativeness of the integral term, the two possible cases $\dot{\gamma}_t = 0$ and $\dot{\gamma}_t > 0$ are separately considered. (i) if $\dot{\gamma}_t = 0$, then $\delta\dot{\gamma} \geq 0$ for (38), and the integral term in (36) is nonnegative if $f_t + \tau \dot{f}_t \geq 0$; (ii) if $\dot{\gamma}_t > 0$, $\delta\dot{\gamma}$ can have any sign, and the integral term in (36) is nonnegative if $f_t + \tau \dot{f}_t = 0$. Thus, summing up, evolution is governed by the Kuhn-Tucker conditions

$$\dot{\gamma}_t \geq 0, \quad f_t + \tau \dot{f}_t \geq 0, \quad (f_t + \tau \dot{f}_t) \dot{\gamma}_t = 0, \tag{39}$$

which represent the flow rule of plasticity. It states that the stress maintains equal to the yield stress when γ evolves.

Equations (27)₁, (39) and the boundary conditions (28) are necessary conditions for a minimum of the total energy $E_{t+\tau}$, i.e., for $(\dot{u}_t, \dot{\gamma}_t)$ to be solution of the following constrained quadratic programming problem

$$(\dot{u}_t, \dot{\gamma}_t) = \operatorname{argmin}\{\Delta E_t(\dot{u}_t, \dot{\gamma}_t), \dot{\gamma}_t \geq 0, \text{ and b.c. on } \dot{u}_t \text{ and } \dot{\gamma}_t\}. \tag{40}$$

In the numerical code, problem (40) is solved by implementing a constrained quadratic programming algorithm.

Analytical Solutions

In this section, the evolution problems described in the previous sections are analytically solved in the simplified special case of evolution from homogeneous configurations. The analytical solutions found in Del Piero et al. (2013) and in Lancioni (2015) for the RI model and in Lancioni et al. (2015a) for the RD model are recalled and commented, since they provide useful criteria for the choice of the plastic energy functionals, which are the free energy ψ_γ in the RD model and the dissipative potential θ in the RI model, whose shape is still unspecified.

At the instant t , we assume that γ_t is homogenous and that, for the RI model, the yield condition (35) is satisfied as an equality, i.e., $\sigma_t = \theta'_t$, since $\xi'_t = 0$ for constant γ . The boundary conditions are

$$\dot{u}_t(0) = 0, \quad \dot{u}_t(l) = l\dot{\varepsilon}_t, \quad \dot{\gamma}_t(0) = \dot{\gamma}_t(l) = 0, \quad (41)$$

where the most interesting case of hard boundary conditions for $\dot{\gamma}_t$ is considered.

RD Model

Under the above simplified assumptions, the evolution equation (31)₁ reduces to a differential equation with constant coefficients, which can be written in the form

$$A\dot{\gamma}_t'' - \left(\psi_\gamma'' + \frac{c}{\delta\varepsilon_t} \dot{\varepsilon}_t \right) \dot{\gamma}_t = - \left(\dot{\sigma}_t + \frac{\sigma^d}{\delta\varepsilon_t} \dot{\varepsilon}_t \right), \quad (42)$$

where $\psi_\gamma'' = \psi_\gamma''(\gamma_t)$ is constant. In order to make explicit the dependence on the deformation rate $\dot{\varepsilon}$, we have used the relation $\tau = \delta\varepsilon/\dot{\varepsilon}$, with $\delta\varepsilon$ the deformation increment in the time step τ . Equation (42) is solved in two steps. In the first step, the solution $\dot{\gamma}_t$ is found as a function of $\dot{\sigma}_t$, and, in the second step, $\dot{\sigma}_t$ is explicitly determined by using the relation

$$\dot{\sigma}_t = E(\dot{\varepsilon}_t - \dot{\gamma}_t), \quad (43)$$

which is obtained by differentiating (7)₁ and using (29). Omitting subscript t for brevity, the solution of the boundary value problem (41) and (42) is

$$\dot{\gamma} = \frac{\dot{\sigma} + \frac{\sigma^d}{\delta\varepsilon} \dot{\varepsilon}}{\psi_\gamma'' + \frac{c\dot{\varepsilon}}{\delta\varepsilon}} \left(1 - \frac{\cosh k_\varepsilon(l/2 - x)}{\cosh k_\varepsilon l/2} \right),$$

$$\dot{\sigma} = E \frac{\psi_\gamma'' + \frac{c\dot{\varepsilon} - \sigma^d}{\delta\varepsilon}}{\psi_\gamma'' + \frac{c\dot{\varepsilon}}{\delta\varepsilon} + E\varphi_1(k_\varepsilon l)} \dot{\varepsilon}, \quad \text{if } \psi_\gamma'' > -\frac{c\dot{\varepsilon}}{\delta\varepsilon},$$

$$\begin{aligned} \dot{\gamma} &= \frac{\dot{\sigma} + \frac{\sigma^d}{\delta\varepsilon}\dot{\varepsilon}}{2A}x(l-x), \quad \dot{\sigma} = E \frac{12A - \frac{\sigma^d l^2}{\delta\varepsilon}}{12A + El^2} \text{ if } \psi''_\gamma = -\frac{c\dot{\varepsilon}}{\delta\varepsilon}, \\ \dot{\gamma} &= \frac{\dot{\sigma} + \frac{\sigma^d}{\delta\varepsilon}\dot{\varepsilon}}{\psi''_\gamma + \frac{c\dot{\varepsilon}}{\delta\varepsilon}} \left(1 - \frac{\cos k_\varepsilon(l/2-x)}{\cos k_\varepsilon l/2} \right), \\ \dot{\sigma} &= E \frac{\psi''_\gamma + \frac{c\dot{\varepsilon}-\sigma^d}{\delta\varepsilon}}{\psi''_\gamma + \frac{c\dot{\varepsilon}}{\delta\varepsilon} + E\varphi_2(k_\varepsilon l)} \dot{\varepsilon}, \text{ if } \psi''_\gamma < -\frac{c\dot{\varepsilon}}{\delta\varepsilon}, \end{aligned} \tag{44}$$

with

$$k_\varepsilon = \sqrt{\frac{1}{A} \left| \psi''_\gamma + \frac{c\dot{\varepsilon}}{\delta\varepsilon} \right|}, \tag{45}$$

and

$$\varphi_1(y) = 1 - \frac{2}{y} \tanh \frac{y}{2}, \quad \varphi_2(y) = 1 - \frac{2}{y} \tan \frac{y}{2}. \tag{46}$$

Functions φ_1 and φ_2 are graphed in Fig. 2.

Formulas (44) refer to three different types of solutions, associated with three different intervals of values of ψ''_γ , as shown by the graphs of Fig. 3a. Hyperbolic solutions characterize the semiaxis $\psi''_\gamma > -\frac{c\dot{\varepsilon}}{\delta\varepsilon}$. In the semiaxis $\psi''_\gamma < -\frac{c\dot{\varepsilon}}{\delta\varepsilon}$, two types of trigonometric solutions distinguish: for $-\frac{4\pi^2 A}{l^2} - \frac{c\dot{\varepsilon}}{\delta\varepsilon} \leq \psi''_\gamma < -\frac{c\dot{\varepsilon}}{\delta\varepsilon}$, the plastic strain rate is positive in any points of the domain, while, for $\psi''_\gamma < -\frac{4\pi^2 A}{l^2} - \frac{c\dot{\varepsilon}}{\delta\varepsilon}$, the plastic strain rate is positive in a central part of the domain and negative in portions close to the endpoints. This latter situation describes a process of strain localization in the center of the domain and plastic strain recovery on the sides. Plastic strain unloading is possible since the plastic energy is assumed to be stored.

Notice that the sign of $\dot{\sigma}$ depends on the deformation rate $\dot{\varepsilon}$. If we suppose that σ^d is so small that it can be neglected, then $\dot{\sigma}$ is positive for $\psi''_\gamma > -(\pi^2 A/l^2 - c\dot{\varepsilon}/\delta\varepsilon)$, and it is negative for $\psi''_\gamma < -(\pi^2 A/l^2 - c\dot{\varepsilon}/\delta\varepsilon)$. In the former case, the evolution regime is stress-hardening, while, in the latter case, it is stress-softening.

RI Model

Since we have assumed that $\sigma_t = \psi'_\gamma$, then $f_t = 0$, and the evolution relations (39) simplify as follows:

$$\dot{\gamma}_t \geq 0, \quad \dot{\sigma}_t \leq \theta''_t \dot{\gamma}_t - \alpha \dot{\gamma}''_t, \quad (\theta''_t \dot{\gamma}_t - \dot{\sigma}_t - \alpha \dot{\gamma}''_t) \dot{\gamma}_t = 0. \tag{47}$$

We introduce the coefficient

$$k_0 = \sqrt{\frac{|\theta''|}{\alpha}}, \quad (48)$$

which is equal to the coefficient (45), with $\dot{\varepsilon} = 0$ and $\psi_\gamma(\gamma) = \theta(\gamma)$. First, we suppose $\theta'' \geq 0$. Integrating (47)₃ over $(0, l)$, we obtain

$$\dot{\sigma} \int_0^l \dot{\gamma} dx = \int_0^l (\theta'' \dot{\gamma}^2 + A \dot{\gamma}'^2) dx > 0, \quad (49)$$

from which $\dot{\sigma} > 0$. If we suppose $\dot{\gamma} = 0$ in some intervals of $(0, l)$, from (47)₂, $\dot{\sigma} \leq 0$, in contradiction with the above result. Thus $\dot{\gamma}$ must be positive almost everywhere in $(0, l)$, and the Khun-Tucker conditions (47) are satisfied if

$$\dot{\gamma}'' - k_0^2 \dot{\gamma} = -\frac{\dot{\sigma}}{A}, \quad \text{for any } x \in (0, l), \quad (50)$$

where index t is omitted. Solutions of (50) satisfying the boundary conditions (41) are found by following the procedure of section “RD Model”: first the solution $\dot{\gamma}_t$ is found in terms of $\dot{\sigma}_t$, and, then, $\dot{\sigma}_t$ is explicitly determined from (43). We obtain

$$\dot{\gamma} = \frac{\dot{\sigma}}{\theta''} \left(1 - \frac{\cosh k_0 \left(\frac{l}{2} - x\right)}{\cosh k_0 \frac{l}{2}} \right), \quad \dot{\sigma} = E \frac{\theta''}{\theta'' + E\varphi_1(k_0 l)} \dot{\varepsilon} > 0, \quad (51)$$

with φ_1 equal to (46)₁. Notice that $\dot{\gamma} > 0$ for any $x \in (0, l)$. Also $\dot{\sigma} > 0$, since numerator and denominator in (51)₂ are positive. It follows that, if $\theta'' > 0$, plastic strain evolves in the whole domain (*full-size* solution) under a regime of stress-hardening. If $\theta'' = 0$, Eq. (50) reduces to $A\dot{\gamma}'' = \dot{\sigma}$, whose solution is

$$\dot{\gamma} = \frac{\dot{\sigma}}{2A} x(l-x), \quad \dot{\sigma} = E \frac{12A}{12A + El^2} \dot{\varepsilon} > 0. \quad (52)$$

Even in this case, $\dot{\gamma}$ is full-size, and the evolution process is stress-hardening. For $\theta'' < 0$, Eq. (47)₃ becomes elliptic

$$\dot{\gamma}'' + k_0^2 \dot{\gamma} = -\frac{\dot{\sigma}}{A}, \quad (53)$$

and it admits the solution

$$\dot{\gamma} = \frac{\dot{\sigma}}{\theta''} \left(1 - \frac{\cos k_0 \left(\frac{l}{2} - x\right)}{\cos k_0 \frac{l}{2}} \right), \quad \dot{\sigma} = E \frac{\theta''}{\theta'' + E\varphi_2(k_0 l)} \dot{\varepsilon}, \quad (54)$$

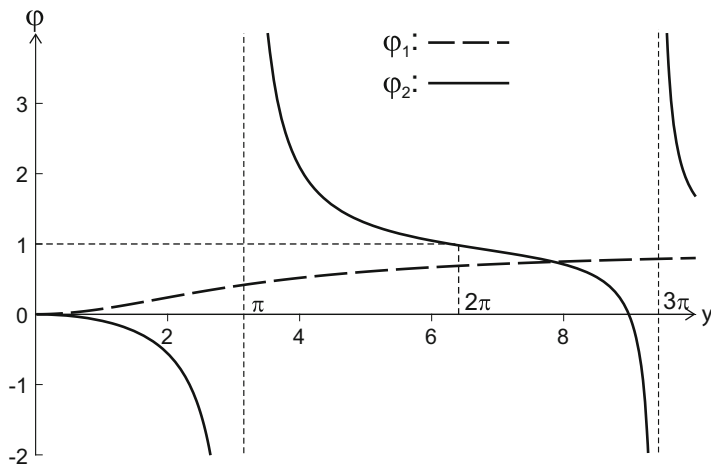


Fig. 2 Graphs of functions $\varphi_1 = \varphi_1(y)$ and $\varphi_2 = \varphi_2(y)$

with φ_2 as in (46)₂. Solution (54) is positive for any $x \in (0, l)$, if $-4\pi^2 A/l^2 \leq \theta'' < 0$. For $\theta'' < -4\pi^2 A/l^2$, the above solution is negative in portions of $(0, l)$ close to the endpoints, in contradiction with the irreversibility condition (33), and, thus, localized solutions must be looked for. Before determining localized solutions, we discuss the sign of $\dot{\sigma}$ in (54)₂. If $-\pi^2 A/l^2 < \theta'' < 0$ (i.e., $0 < k_0 l < \pi$), then $\dot{\sigma} > 0$. Indeed, being $\varphi_2 < 0$ (see Fig. 2), both numerator and denominator in (54)₂ are negative, and the evolution regime is stress-hardening. For $-4\pi^2 A/l^2 \leq \theta'' < -\pi^2 A/l^2$ ($\pi < k_0 l \leq 2\pi$), φ_2 is positive, and $\dot{\sigma} \leq 0$, i.e., stress-softening regime, if

$$\theta'' + E\varphi_2(k_0 l) > 0. \tag{55}$$

It was proven in Del Piero et al. (2013) that inequality (55) is a sufficient condition for stability of solution (54). Since $\varphi_2 \geq 1$, if $-4\pi^2 A/l^2 \leq \theta'' < -\pi^2 A/l^2$, the above inequality is satisfied if $E > 4\pi^2 A/l^2$, which is a condition usually met by real applications.

Now, we consider the case $\theta'' < -4\pi^2 A/l^2$, and we look for localized solutions. We assume that $\dot{\gamma}$ localizes in a portion $(0, \hat{l})$, adjacent to the bar left endpoint, and the supplementary continuity conditions $\dot{\gamma}(\hat{l}) = \dot{\gamma}'(\hat{l}) = 0$ are assigned. Equation (53) is solved within the localization zone $(0, \hat{l})$, and the inequality $\dot{\sigma} \leq 0$ is satisfied outside it, where $\dot{\gamma} = 0$. The resulting solution is

$$\begin{cases} \dot{\gamma} = \frac{\dot{\sigma}}{\theta''} (1 - \cos k_0 x), & \text{for } 0 < x \leq \hat{l}, \text{ with } \hat{l} = \frac{2\pi}{k_0}, \\ \dot{\gamma} = 0, & \text{for } \hat{l} < x < l, \end{cases} \tag{56}$$

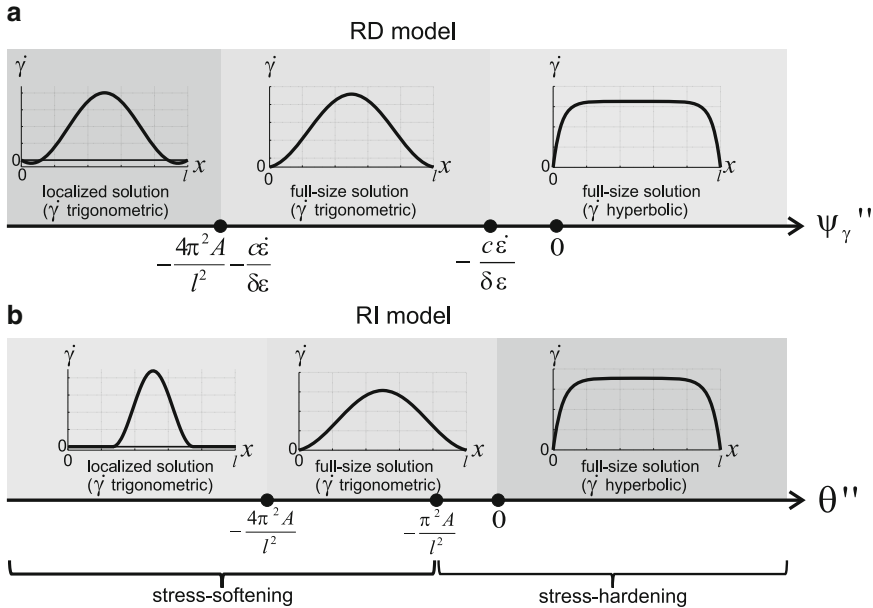


Fig. 3 Scheme of analytical solutions in the case of RD model (a), and RI model (b)

where the length \hat{l} is determined from the supplementary condition $\dot{\gamma}'(\hat{l}) = 0$, and

$$\dot{\sigma} = E \frac{\theta''}{\theta'' + 2\pi \frac{E}{k_0 l}} \dot{\epsilon}. \tag{57}$$

Notice that any translation of the localized solution (56) is also a solution. The stress rate $\dot{\sigma}$ is negative (stress-softening regime), if

$$\theta'' + 2\pi \frac{E}{k_0 l} > 0, \tag{58}$$

which, as proved in Del Piero et al. (2013), is a sufficient condition for stability of (56), analogously to the stability relation (55). When θ'' approaches $-2\pi E / (k_0 l)$ from above, then $\dot{\sigma} \rightarrow -\infty$, and $\dot{\gamma} \rightarrow \infty$ in a small region of the domain. The total energy associated with this configuration drops to minus infinity. This situation describes the occurrence of fracture. Detailed analyses of fractured configurations are proposed in Del Piero et al. (2013), Del Piero (2013), and Lancioni (2015).

Figure 3b schematizes the different solutions described above, which depend on the value of θ'' .

Numerical Results

The RD and RI evolution problems, governed by Eqs. (28) and (31) and the constrained quadratic programming problem (40), respectively, are implemented numerically by finite elements. In each one-dimensional finite element, the plastic strain γ is approximated by a linear shape function, depending on two nodal unknowns defined at the element endpoints, while the displacement u is approximated by a quadratic shape function, where three nodal variables are defined at the element endpoints and midpoint. The solution of the evolution problems is refined at each time step by means of a Newton-Raphson iterative scheme in the RD formulation and by implementing a sequential quadratic programming algorithm in the RI model.

In the following, the numerical tests presented in Lancioni et al. (2015a) are discussed, where two distinct non-convex plastic energies are assigned in order to reproduce two different processes: plastic shear slip patterning in metals (section “Plastic Slip Patterning”) and necking in tensile steel bars (section “Necking in Tensile Steel Bars”).

Plastic Slip Patterning

The problem of formation and evolution of microstructures in metallic materials is addressed in the simplified one-dimensional geometrical scheme proposed in Fig. 1b. A semi-infinite layer with thickness l is subjected to a shear deformation ε . The evolution of the plastic shear strain γ through the thickness is analyzed in both RD and RI frameworks, by first assuming the soft boundary conditions (4)₁ for γ and then the hard boundary conditions (4)₂. The layer thickness is $l = 1$ mm. The Young modulus E used in the above theoretical sections is replaced by the shear modulus $G = 78.947$ GPa, which corresponds to the Young modulus $E = 210$ GPa and Poisson’s ratio $\nu = 0.33$ of steel. As in Yalçinkaya et al. (2011), we fix $c = 7$ MPa/s and $A = 147.29$ N. Coefficient A is related to the internal length scale by the formula $A = ER^2/(16(1 - \nu^2))$ proposed in Bayley et al (2006), where the length scale R represents the radius of the dislocation domain contributing to the internal stress field. In this example $R = 0.1$ mm. A Landau-Devonshire type of potential $f_1 = f_1(\gamma)$ is assigned to the plastic energy densities ψ_γ (RD model) and θ (RI model), whose expression is

$$f_1(\gamma) = 1.525 \times 10^8 \gamma^4 - 5.2 \times 10^6 \gamma^3 + 5.25 \times 10^4 \gamma^2 \text{ MPa.} \quad (59)$$

The graphs of f_1 and of its first derivative are represented in Fig. 4. Values of binodal and spinodal points are indicated in the graphs, as well the stress corresponding to the Maxwell line.

When soft boundary conditions are applied, the response of the layer to the shear deformation ε is described by the stress versus strain curves of Fig. 5. The RI model

Fig. 4 Double-well plastic energy density (a) and its derivative (b)

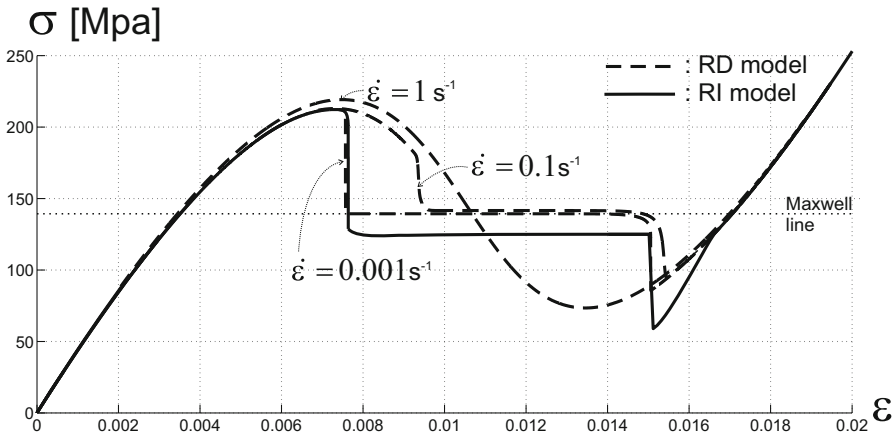
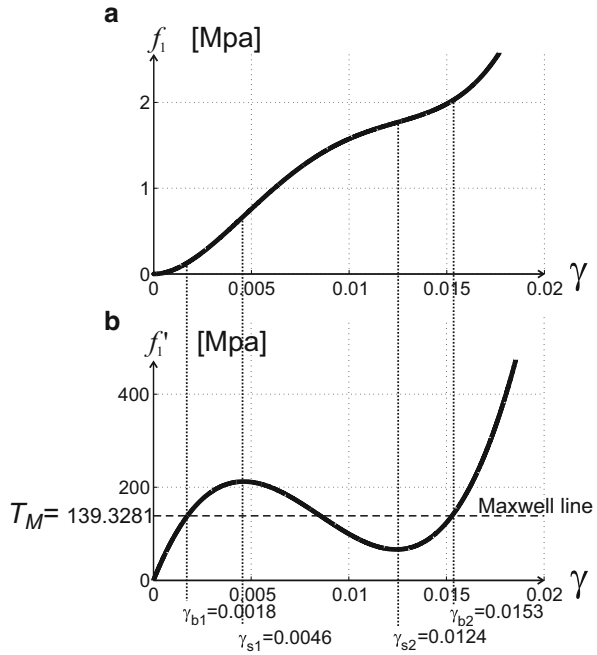


Fig. 5 Response curves in case of soft boundary conditions. Deformation rates $\dot{\epsilon} = 0.001, 0.1, 1 \text{ s}^{-1}$ are considered in the RD problem

provides the solid line curve, while the dashed line curves are obtained from the RD model, by assigning three different deformation rates, i.e., $\dot{\epsilon} = 0.001, 0.1,$ and 1 s^{-1} . Notice that the response curves of the RD model get closer to the curve of the RI model, as $\dot{\epsilon}$ decreases, because the effect of the viscous stress reduces. Three branches characterizes the response curves: (i) an initial hardening branch;

(ii) an intermediate softening phase, in the RI case and in RD cases with low rates, exhibiting sharp initial and final drops, and central plateaus; and (iii) a final hardening curve. These three steps of the curves are associated with different plastic strain regimes, as described in Fig. 6, which show different profiles of γ at different values of the imposed deformation.

In the initial hardening phase, the plastic strain grows homogeneously in the whole domain, which is common to all the four simulations. This phase ends when the plastic strain (plastic slip in this one-dimensional case) γ_{1s} , corresponding to the first spinodal point, is reached. From then on, different evolutions are observed. The RI simulation and the RD simulations with sufficiently low deformations rates ($\dot{\epsilon} = 0.001, 0.1 \text{ s}^{-1}$) exhibit strain localization on the right side of the domain, corresponding to the stress dropping down in the response curves. This stress drop occurs and is postponed as the deformation rate increases. The viscous stress contribution taken into account in the RD model produces a delay of the localization process. This can be clearly noticed by comparing the profiles of γ in Fig. 6b, c corresponding to the localization stage. Afterward, the plastic band moves from the right to the left. The plastic wave propagation corresponds to the stress plateau of the $\sigma - \epsilon$ curves. The stress value of the plateau lowers as $\dot{\epsilon}$ decreases, approaching to the Maxwell line. Differently, in the RI case, the plateau is below the Maxwell line.

The evolutions predicted by the two models in this intermediate phase of heterogeneous strain evolution exhibit a further crucial difference pointed out in the following. In the RD model, plastic strain localizes till the saturation value γ_{b2} , corresponding to the second binodal point, is reached in the right part of the body, while at the left side γ rapidly reduces to the value γ_{b1} of the first binodal point. Although the plastic wave front for the RI model has a similar width, the wave peak has a value slightly lower than γ_{b2} , and the strain downstream remains constant at the value γ_{s1} of the first spinodal point, the value at which localization is initiated. This is explained by the fact that the plastic energy is totally dissipated in the RI model, and thus γ can only grow. On the other hand, the energy ψ_γ of the RD model is stored, and therefore γ can be partially recovered, which justifies the reduction of γ outside the localization zone. Plastic energy recovery is also evident at the end of the slip patterning evolution, when the deformation becomes homogeneous. The evolution step from $\epsilon = 0.0150$ to $\epsilon = 0.0151$ results in a significant reduction in γ over the entire domain. The change to homogeneous strain is accompanied by a further stress drop, as shown by the curves of Fig. 5. This strain patterning is totally lost by the RD model when sufficiently large deformation rates are considered. For $\dot{\epsilon} = 1 \text{ s}^{-1}$, evolution is always homogeneous, as shown in Fig. 6d, even if the stress-strain curve exhibits a softening branch.

Constitutive response curves in the case of hard boundary conditions are plotted in Fig. 7. Their shapes are similar in type to those of the previous case of soft boundary conditions, that is, they present two hardening branches and an interposed softening stage. In this case they present a smoother shape. However, also the curves of Fig. 5 are always continuous, and the drops, which apparently look discontinuous, are smooth quickly decreasing curves.

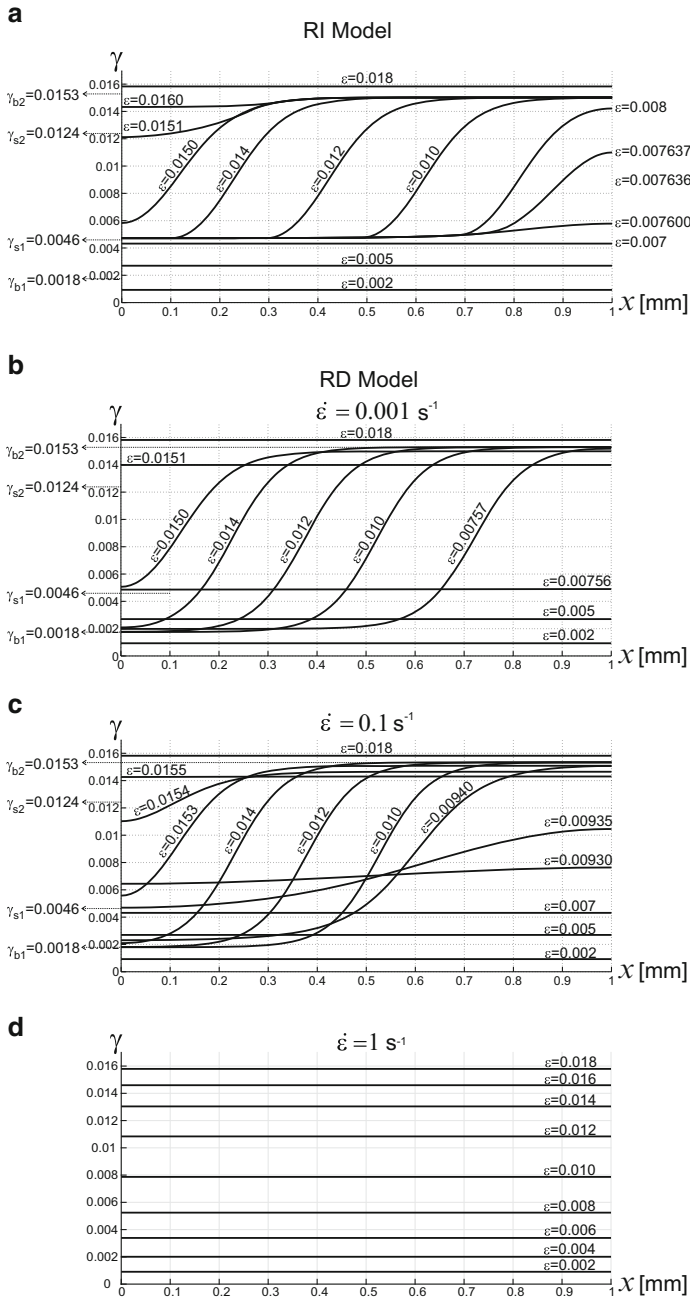


Fig. 6 Profiles of γ at different values of $\dot{\epsilon}$ in case of soft boundary conditions. (a) RI model, (b) RD model with $\dot{\epsilon} = 0.001 \text{ s}^{-1}$, (c) RD model with $\dot{\epsilon} = 0.1 \text{ s}^{-1}$, (d) RD model with $\dot{\epsilon} = 1 \text{ s}^{-1}$

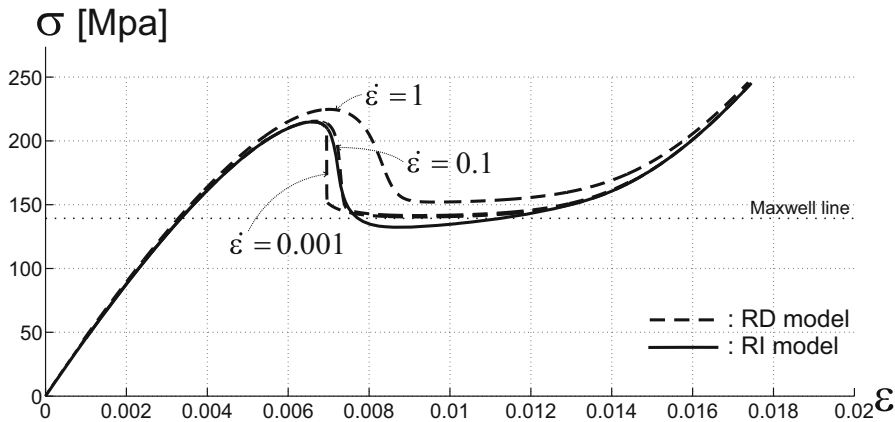


Fig. 7 Response curves in case of hard boundary conditions. Deformation rates $\dot{\epsilon} = 0.001, 0.1, 1 \text{ s}^{-1}$ are considered in the RD problem

In Fig. 8, snapshots of γ are taken at deformation increment $\Delta\epsilon = 10^{-3}$. The evolution process that they describe is clearly composed by four phases: (1) an initial hardening phase in which deformation is homogeneous, except in the boundary layers developed at each end; (2) a softening process of strain localization in the middle of the domain; (3) a subsequent stress plateau characterized by plastic slip spreading toward the boundaries; and (4) a final hardening branch, where the deformation field evolves in a similar way to the initial hardening phase. If the evolution profiles of Fig. 8b, c are compared, it can be noticed that the increase of the deformation rate produces a delay of the localization process, which only partially takes place. The size of the zone where γ initially localizes becomes larger and larger as $\dot{\epsilon}$ increases.

Necking in Tensile Steel Bars

Now, we consider an homogeneous bar of length $l = 140 \text{ mm}$, clamped at the left endpoint and subjected to the tensile displacement ϵl at the right endpoint (see Fig. 1b). The bar is made of steel, with Young’s Modulus: $E = 210 \text{ GPa}$. For the plastic energy (ψ_γ in the RD model and θ in the RI model), we consider the piecewise cubic function considered in Lancioni (2015). This function, that we indicate with $f_2 = f_2(\gamma)$, is graphed in Fig. 9, with its first and second derivatives, and its analytical expression is

$$f_2(\gamma) = \begin{cases} c_1\gamma + \frac{1}{2}c_2\gamma^2 \left(1 - \frac{\gamma}{3\gamma_1}\right), & \text{if } 0 < \gamma \leq \gamma_1, \\ c_3\gamma + c_4(\gamma - \gamma_1) + c_5(\gamma - \gamma_1)^3, & \text{if } \gamma_1 < \gamma \leq \gamma_2, \end{cases} \quad (60)$$

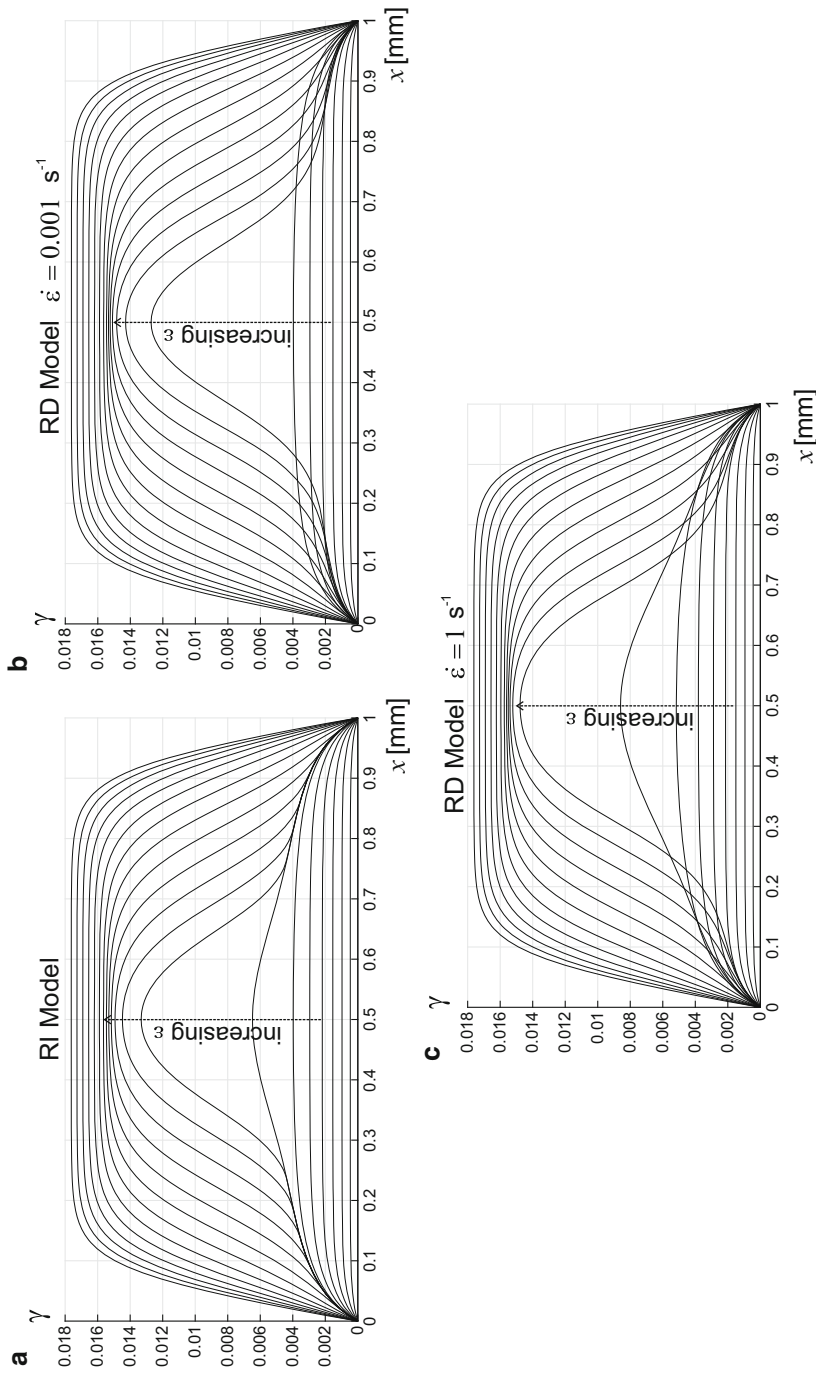


Fig. 8 Snapshots of γ taken at increment of ϵ equal to 10^{-3} . (a) RI model, (b) RD model with $\dot{\epsilon} = 0.001 \text{ s}^{-1}$, (c) RD model with $\dot{\epsilon} = 1 \text{ s}^{-1}$

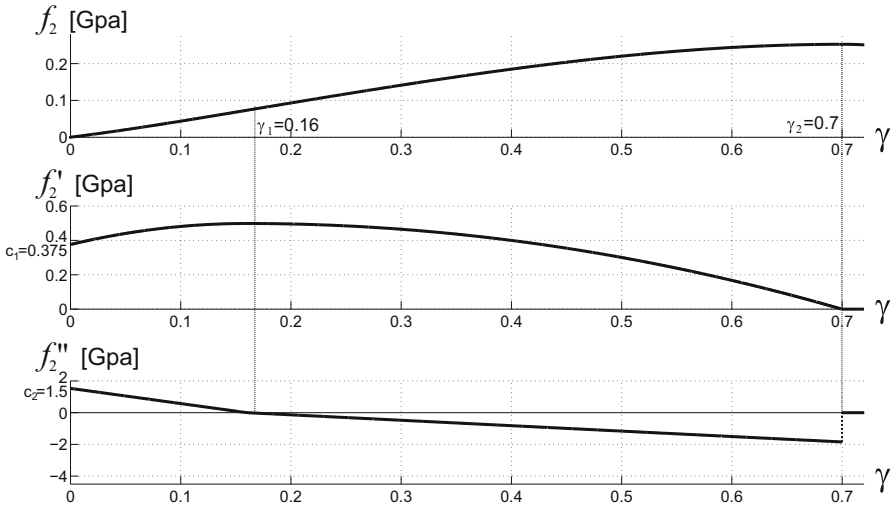


Fig. 9 Plastic energy density of the steel bar and its first and second derivatives

where $c_3 = c_1\gamma_1 + c_2\gamma_1^2/6$, $c_4 = c_1 + c_2\gamma_1/2$, and $c_5 = -(2c_1 + c_2\gamma_1)/(6(\gamma_2 - \gamma_1)^2)$ to guarantee the continuity of θ and θ' at γ_1 and the condition $f_2' \geq 0$. Coefficients c_1, c_2, γ_1 , and γ_2 are related to easily measurable experimental quantities through the relations

$$c_1 = \sigma^c, \quad c_2 = \frac{2}{\gamma_1}(\sigma^m - \sigma^c), \quad \gamma_1 = \varepsilon^m - \frac{\sigma^m}{E}, \quad \gamma_2 = \frac{2(d_0 - d_{break})}{d_0}, \quad (61)$$

where σ^c is the yield stress, σ^m is the peak stress, ε^m is the corresponding strain, d_0 are the initial diameters of the bar cross section, and d_{break} is the diameter of the broken cross section. We assign the following values obtained from tensile tests

$$\sigma^c = 0.376 \text{ GPa}, \quad \sigma^m = 0.496 \text{ GPa}, \quad \varepsilon^m = 0.15, \quad d_0 = 10 \text{ mm}, \quad d_{brak} = 6.5 \text{ mm}. \quad (62)$$

Deduction and detailed description of formulas (61) are given in Lancioni (2015). The energy functional f_2 is convex in the interval $0 \leq \gamma < \gamma_1$, concave for $\gamma_1 \leq \gamma < \gamma_2$, and constant for $\gamma > \gamma_2$. For $\gamma > \gamma_2$, the bar deforms without spending plastic energy, and this corresponds to complete breaking. Thus, the simulations presented in the following are interrupted when γ reaches the breaking value $\gamma_2 = 0.7$. For the nonlocal parameter, we assume $A = 2 \text{ kN}$, as in Lancioni (2015), and, for the viscous coefficient of the RD model, we fix $c = 0.015 \text{ GPa s}^{-1}$ as in Lancioni et al. (2015a).

Macroscopic constitutive response ($\sigma - \varepsilon$) curves are compared in Fig. 10. For the RD model, the four different deformation rates $\dot{\varepsilon} = 10^{-3}, 10^{-1}, 1$, and 10 s^{-1} are considered. In the enlargement on the left side of Fig. 10, the experimental curve is

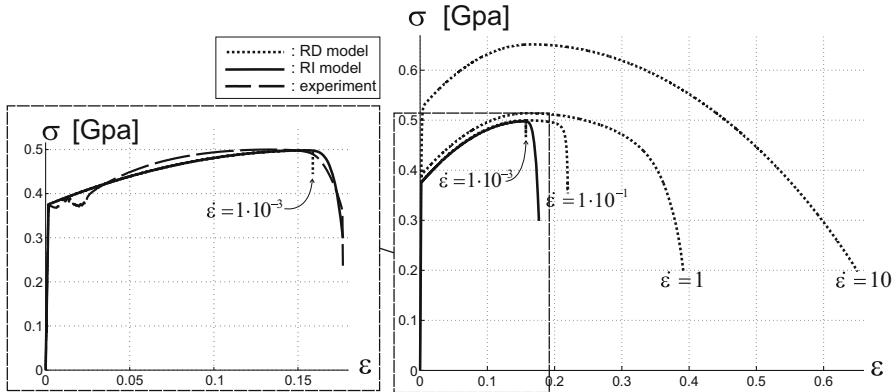


Fig. 10 Response curves of a tensile steel bar

also plotted (dashed line). The simulations predict three phases: an initial perfectly elastic phase, which interrupts when σ reaches the yield value σ^c , a hardening phase, and a final softening phase. In Lancioni (2015), a more sophisticated plastic energy was considered to reproduce the yielding plateau observed at the onset of the inelastic regime (see the experimental curve in the left Fig. 10).

For sufficiently low deformation rates ($\dot{\varepsilon} = 10^{-3}$ and 10^{-1} s^{-1}), the RI and RD models give practically the same hardening branches, with equal yield and peak stresses, but they differ in predicting the softening curves. The branch of the RI model is very close to the experimental curve, while those of the RD model strongly depends on $\dot{\varepsilon}$, getting longer as the deformation rate increases. Since $\dot{\varepsilon} = 1 \cdot 10^{-3}$ is a very small rate, the corresponding curve is a good approximation of the response curve in the limit static case $\dot{\varepsilon} = 0$. The discrepancies between this curve and that of the RI model clearly prove that the RI model is not the limit case of the RD model for $\dot{\varepsilon} \rightarrow 0$, although the RD model gives results close to those of the RI model when small deformation rates are considered. For large values of the deformation rate ($\dot{\varepsilon} = 1$ and 10 s^{-1}), the response curves of the RD model largely deviates from the curve of the RI model. As $\dot{\varepsilon}$ grows, the yield and peak stresses increases, and the softening branches extends. Also the plastic strain evolution is completely different than that of the RI case, as described in the following.

The evolution of γ is described in Fig. 11, where profiles of γ at different values of ε are plotted for the RI case and for the RD one as well with $\dot{\varepsilon} = 10^{-3}$, 1 , 10 s^{-1} . Results of the RI model and RD model with $\dot{\varepsilon} = 10^{-3} \text{ s}^{-1}$ are comparable. In both the cases, the hardening phase is associated with the evolution of homogeneous plastic strains (with the exception of small boundary layers where γ decreases to zero), and the softening phase is characterized by a progressive localization of γ in smaller and smaller portions in the center of the bar, up to the final fracture, occurring when γ reaches the breaking value $\gamma_2 = 0.7$ in the bar midpoint. Both models describe fracture as the ending stage of a strain localization process; however, they differ in predicting the evolution of γ outside the localization zone:

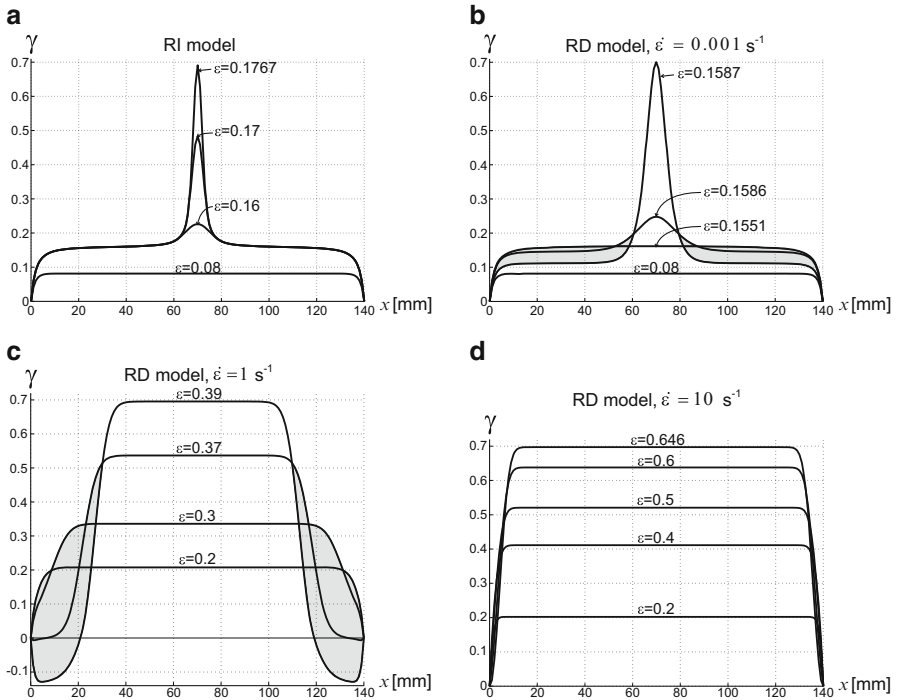


Fig. 11 Profiles of γ at different values of $\dot{\epsilon}$. (a) RI model, (b) RD model with $\dot{\epsilon} = 0.001 \text{ s}^{-1}$, (c) RD model with $\dot{\epsilon} = 1 \text{ s}^{-1}$, (d) RD model with $\dot{\epsilon} = 10 \text{ s}^{-1}$

γ maintains constant in the RI case, and only the elastic deformation ϵ reduces, since $\delta\epsilon = \delta\sigma/E < 0$ (elastic unloading), while both γ and ϵ^e reduce in the RD process (elastoplastic unloading). Gray areas in Fig. 11b highlight the zones where γ is recovered. The partial strain recovery outside the localization portion makes the softening branch of the response curve much steeper than that of the RI curve. The evolution of γ is very different if large $\dot{\epsilon}$ are applied, as illustrated in Fig. 11c, d. Since the plastic strain has no time to localize, the limit value $\gamma = 0.7$ that leads to failure is reached in large zones in the center of the bar. On the contrary, plastic strain is recovered in small zones close to the endpoints. For $\dot{\epsilon} = 1 \text{ s}^{-1}$, plastic recovery leads to local compressive strain states in the final stages of the evolution (see the gray regions in Fig. 11c).

We notice that the above results agree with the analytical solutions found in sections “RD Model” and “RI Model.” The analytical results basically state that solutions are full-size if the plastic energy is convex and localized if it is concave. Accordingly, numerical simulations predict an initial full-size evolution of γ , spreading in the whole bar. The initial hardening phase terminates when γ reach the value $\gamma_1 = 0.16$, at which the plastic energy changes from concave to convex. From this point on, the bar experiences localization. The fact that γ localizes on

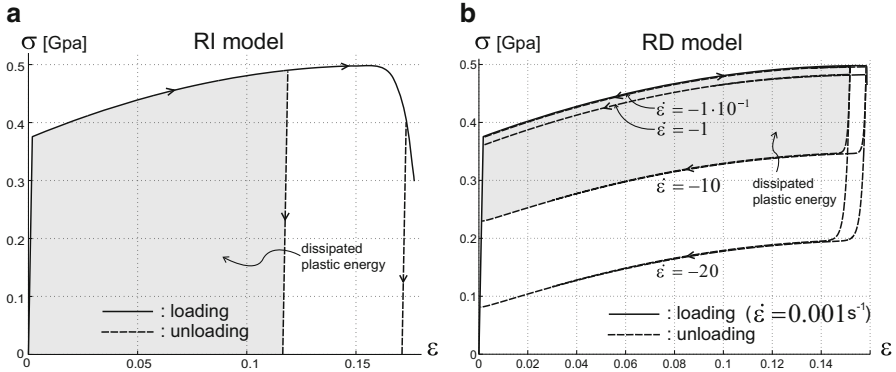


Fig. 12 Loading (solid line) and unloading (dashed line) response curves. (a) RI model, (b) RD model with different unloading plastic strains

smaller and smaller portions is also expected by the analytical results, according to which the size of the localization zone is inversely proportional to $|\psi_\gamma''|$ (see solution (44)₃) or to $|\theta''|$ (solution (56)). In the plastic energy (60) implemented in the simulations, $|f_2''|$ is increasing with γ .

The bar response at unloading is analyzed in Fig. 12. For the RI model, unloading is always elastic, since γ does not reduce. The dissipated plastic energy is represented by the area below the response curve (gray area in Fig. 12a). For the RD model, unloading depends on the deformation rate. For small deformation rates, the unloading curves practically coincide with the loading curve, with no plastic dissipation, but, for larger deformation rates, hysteretic loops are obtained, and the area of the resulting closed curves represents the dissipated viscous energy. These areas are larger and larger as $|\dot{\epsilon}|$ increases. The unloading process given by the RD model exhibits a first elastic unloading, followed by a second elastoplastic unloading, which are clearly recognizable by the different slope of the corresponding branches. The different responses at unloading described by the two models constitute a further distinguishing feature of the two proposed formulations.

Conclusions

In this chapter, two nonlocal plasticity models based on non-convex plastic potentials are addressed, by following a variational procedure which is thermodynamically consistent. Their ability in capturing the evolution of heterogeneous plastic patterning is analyzed in a simplified one-dimensional mathematical setting. The first model is rate-dependent, and it accounts for partially recoverable plastic strains. The second one is rate-independent, and it assumes that plastic strains are totally dissipative. Both plasticity models are enhanced by a gradient energy contribution, which introduces a length parameter into the models allowing studies at different

length scales. Non-convex plastic potentials are assumed that lead to heterogeneous strain distributions and to processes of plastic localization. The instability due to the presence of a non-convex plastic energy is stabilized by the gradient energy term.

The RD and RI problems governing the strain evolution are solved analytically, for a simplified case. Analytical solutions give useful information on the shape that should be assigned to the plastic energy in order to predict specific behaviors, such as stress-hardening with a diffused plastic strain evolution, or stress-softening with strain localization. Both formulations are incorporated into finite element solution procedure and two different problems are solved, each one characterized by a specific non-convex plastic energy and by a specific length scale.

Numerical results have pointed out clearly the similarities and differences of the models, in some cases confirming the analytical evidences. At low deformation rates, the results of the RD model approach to those of the RI formulation, because of the effect of the viscous stress contribution. Differently, at high deformation rates, viscosity produces a delay of the strain localization, which, for particularly large deformation rates, is even missed. Consequently, results largely deviate from those of the RI model.

Differences that are independent on the deformation rate have been also observed, which are related to the different assumptions on the dissipative nature of the plastic energy made within each formulation. Two main differences deserve to be mentioned. (1) The different strain recovery processes that activate outside the strain localization zone: in the RI case, only the elastic strain reduces, while, in the RD case, both elastic and plastic strains are recovered. (2) The different responses at unloading: purely elastic unloading is reproduced by the RI model, with plastic strains cumulated in the body, and elastoplastic unloading is described by the RD case. In this latter case, hysteretic loops are observed and no residual strains cumulate in the body.

References

- E.C. Aifantis, On the microstructural origin of certain inelastic models. *J. Eng. Mater. Technol.* **106**, 326–330 (1984)
- C.J. Bayley, W.A.M. Brekelmans, M.G.D. Geers, A comparison of dislocation induced back stress formulations in strain gradient crystal plasticity. *Int. J. Solids Struct.* **43**, 7268–7286 (2006)
- Z.P. Bazant, M. Jirásek, Nonlocal integral formulations of plasticity and damage: survey of progress. *J. Eng. Mech.* **128**, 1119–1149 (2002)
- G. Del Piero, A variational approach to fracture and other inelastic phenomena. *J. Elast.* **112**(1), 3–77 (2013)
- G. Del Piero, G. Lancioni, R. March, A diffuse cohesive energy approach to fracture and plasticity: the one-dimensional case. *J. Mech. Mater. Struct.* **8**(2–4), 109–151 (2013)
- P. Gudmundson, A unified treatment of strain gradient plasticity. *J. Mech. Phys. Solids* **52**, 1379–1406 (2004)
- M. Gurtin, L. Anand, Thermodynamics applied to gradient theories involving the accumulated plastic strain: the theories of Aifantis and Fleck & Hutchinson and their generalization. *J. Mech. Phys. Solids* **57**, 405–421 (2009)

- M.E. Gurtin, E. Fried, L. Anand, *The Mechanics and Thermodynamics of Continua* (Cambridge University Press, New York, 2010)
- D. Hull, Orientation and temperature dependence of plastic deformation processes in 3.25 percent silicon iron. *Proc. R. Soc. A* **274**, 5–24 (1963)
- M. Jirásek, S. Rolshoven, Localization properties of strain-softening gradient plasticity models. Part II. Theories with gradients of internal variables. *Int. J. Solids Struct.* **46**, 2239–2254 (2009)
- B. Klusemann, T. Yalçinkaya, Plastic deformation induced microstructure evolution through gradient enhanced crystal plasticity based on a non-convex Helmholtz energy. *Int. J. Plast.* **48**, 168–188 (2013)
- B. Klusemann, T. Yalçinkaya, M.G.D. Geers, B. Svendsen, Application of non-convex rate dependent gradient plasticity to the modeling and simulation of inelastic microstructure development and inhomogeneous material behavior. *Comput. Mater. Sci.* **80**, 51–60 (2013)
- G. Lancioni, Modeling the response of tensile steel bars by means of incremental energy minimization. *J. Elast.* **121**, 25–54 (2015)
- G. Lancioni, T. Yalcinkaya, A. Cocks, Energy based non-local plasticity models for deformation patterning, localization and fracture. *Proc. R. Soc. A* **471**, 20150275:1–20150275:23 (2015a)
- G. Lancioni, G. Zitti, T. Yalcinkaya, Rate-independent deformation patterning in crystal plasticity. *Key Eng. Mater.* **651–653**, 944–949 (2015b)
- A. Mielke, A mathematical framework for generalized standard materials in the rate-independent case, in *Multifield Problems in Solid and Fluid Mechanics*, ed. by R. Helmig, A. Mielke, B. Wohlmuth. *Lecture Notes in Applied and Computational Mechanics*, vol. 28 (Springer, Berlin/Heidelberg, 2006), pp. 399–428
- T. Yalçinkaya, Microstructure evolution in crystal plasticity: strain path effects and dislocation slip patterning. Ph.D. thesis, Eindhoven University of Technology (2011)
- T. Yalçinkaya, G. Lancioni, Energy-based modeling of localization and necking in plasticity. *Procedia Mater. Sci.* **3**, 1618–1625 (2014)
- T. Yalçinkaya, W.A.M. Brekelmans, M.G.D. Geers, Deformation patterning driven by rate dependent nonconvex strain gradient plasticity. *J. Mech. Phys. Solids* **59**, 1–17 (2011)
- T. Yalçinkaya, W.A.M. Brekelmans, M.G.D. Geers, Non-convex rate dependent strain gradient crystal plasticity and deformation patterning. *Int. J. Solids Struct.* **49**, 2625–2636 (2012)



Strain Gradient Crystal Plasticity: Thermodynamics and Implementation

28

Tuncay Yalçinkaya

Contents

Introduction	1002
Plastic Slip-Based Strain Gradient Crystal Plasticity	1004
Rate Variational Formulation of Strain Gradient Crystal Plasticity	1013
Finite Element Solution Procedure of Strain Gradient Crystal Plasticity Framework	1016
Simulation of Polycrystalline Behavior	1025
Conclusion and Outlook	1029
References	1030

Abstract

This chapter studies the thermodynamical consistency and the finite element implementation aspects of a rate-dependent nonlocal (strain gradient) crystal plasticity model, which is used to address the modeling of the size-dependent behavior of polycrystalline metallic materials. The possibilities and required updates for the simulation of dislocation microstructure evolution, grain boundary-dislocation interaction mechanisms, and localization leading to necking and fracture phenomena are shortly discussed as well. The development of the model is conducted in terms of the displacement and the plastic slip, where the coupled fields are updated incrementally through finite element method. Numerical examples illustrate the size effect predictions in polycrystalline materials through Voronoi tessellation.

T. Yalçinkaya (✉)

Aerospace Engineering Program, Middle East Technical University Northern Cyprus Campus,
Guzelyurt, Mersin, Turkey

Department of Aerospace Engineering, Middle East Technical University, Ankara, Turkey
e-mail: yalcinka@metu.edu.tr

© Springer Nature Switzerland AG 2019

G. Z. Voyiadjis (ed.), *Handbook of Nonlocal Continuum Mechanics for Materials and Structures*, https://doi.org/10.1007/978-3-319-58729-5_2

1001

Keywords

Size effect · Strain gradient · Crystal plasticity · Finite element · Polycrystalline plasticity · Thermodynamics

Introduction

Strain gradient frameworks have been the most popular plasticity models for the last three decades, since the early work Aifantis (1984) due to their intrinsic capability to capture the size-dependent behavior of metallic materials through the incorporated length scale parameter which is absent in the classical models that are local and has no reference to the microstructural characteristic lengths. There is a large body of literature illustrating the size effect in the different materials under different loading conditions such as torsion (see, e.g., Fleck et al. 1994; Aifantis 1999), bending (see, e.g., Aifantis 1999; Stölken and Evans 1998; Haque and Saif 2003), indentation (see, e.g., Nix and Gao 1998; Swadener et al. 2002), and compression experiments such as Wang et al. (2006) and Volkert and Lilleodden (2006).

There have been developments in both phenomenological-type isotropic strain gradient plasticity models and physics-based strain gradient crystal plasticity frameworks in order to model the size-dependent response of materials and the evolution of the inhomogeneous strain distribution at the interfaces and boundaries. All of them incorporate a length scale, to capture the size effect, however there is no unified structure of such nonlocal frameworks. Depending on the chosen internal state variables, the global degrees of freedom, the way the strain gradients are incorporated, and the thermodynamical work conjugate nature of the models, there have been various classifications in the literature.

Considering the phenomenological models, Hutchinson (2012) mentions that it has not been a simple matter to obtain a sound extension of the classical J2 flow theory of plasticity that incorporates a dependence on plastic strain gradients. Two classes of J2-type extensions have been proposed: one with increments in higher-order stresses related to increments of strain gradients and the other characterized by the higher-order stresses themselves expressed in terms of increments of strain gradients. The theory in Mühlhaus and Aifantis (1991) and Fleck and Hutchinson (2001) are in the first class, and these do not always satisfy the thermodynamic consistency. The other class includes Gudmundson (2004) and Gurtin and Anand (2009) which are thermodynamically consistent; however they have the physical deficiency that the higher-order stress quantities can change discontinuously for bodies subject to arbitrarily small load changes. Hutchinson (2012) presents a sound phenomenological extension of the rate-independent theory of the first class with a modification of the Fleck-Hutchinson model ensuring its thermodynamic integrity. The problems including nonproportional loading have recently been solved using the updated thermodynamically consistent model in Fleck et al. (2014). Reddy (2011a) has recently studied the flow rules of rate-independent gradient plasticity models and analyzed the uniqueness and the existence of solutions. In addition to these higher-order theories, lower-order theories have been developed based on

the plastic strain gradient-dependent work-hardening rules such as Chen and Wang (2000) and Huang et al. (2004). The recent work by Panteghini and Bardella (2016) presents a detailed finite element implementation of isotropic gradient plasticity frameworks.

Initiated by Taylor (1938) crystal plasticity frameworks have been the main physically based approach in plasticity to address the intrinsic anisotropy and the localization phenomena in metals. It offers a kinematically representative description of first-order crystallographic phenomena (slip planes and directions, elastic anisotropy) and based on the statistical representation of the kinetics of groups of dislocations via flow and hardening rules at the individual slip system level. Even though the main application of such models is the small-scale engineering, the conventional ones do not include a length scale to address the realistic size-dependent behavior. Moreover, slip band formation, plastic slip microstructure evolution, and the effect of the grain boundaries are the limiting features of these models. The extension of crystal plasticity models to capture such effects will be discussed in the current and upcoming chapters.

Since the crystal plasticity extension of Fleck and Hutchinson (1997) to illustrate the grain size effect on strength (Shu and Fleck 1999), various strain gradient crystal plasticity models have been proposed. One of the simplest ways to incorporate the length scale into the constitutive frameworks has been the enhancement of the work-hardening laws with strain gradients such as lattice incompatibility-based models (see, e.g., Acharya and Bassani 2000; Bassani 2001), mechanism-based strain gradient crystal plasticity models (see, e.g., Han et al. 2005a,b), and others such as Ohashi (2005), Dunne et al. (2007), and Liang and Dunne (2009). In such models plastic strain gradients correspond to the geometrically necessary dislocation (GND) densities (Ashby 1970), and they contribute to the hardening behavior in addition to the statistically stored dislocation (SSD) densities. This type of lower-order crystal plasticity enhancement does not bring any additional boundary condition to the conventional boundary value problem formulation.

The other approach is the higher-order extension of crystal plasticity frameworks (see Gurtin 2000, 2002; Yefimov et al. 2004; Evers et al. 2004; Arsenlis et al. 2004; Bayley et al. 2006; Ma et al. 2006; Borg 2007; Geers et al. 2007; Levkovitch and Svendsen 2006; Kuroda and Tvergaard 2008; Reddy 2011b; Yalcinkaya 2011; Yalcinkaya et al. 2012; Klusemann and Yalçinkaya 2013; Klusemann et al. 2013) that would allow the treatment of plastic slip/dislocation density boundary conditions, and it is the main focus of the present paper. The higher-order theories can further be subclassified into work-conjugate and non-work-conjugate types (see Kuroda and Tvergaard 2008). In the former one, the plastic slip gradients are accompanied by the thermodynamically conjugate higher-order stress terms in the virtual work expression, and there exists an external contribution of work arising from a corresponding higher-order surface traction (see, e.g., Gurtin 2002; Borg 2007; Yalcinkaya et al. 2012). This unconventional virtual work expression leads to slip or higher-order traction at the boundaries which would allow us to solve problems where strong barriers of a material interface are impenetrable for dislocations. In the non-work-conjugate type of works (such as Yefimov et al. 2004;

Evers et al. 2004; Arsenlis et al. 2004; Bayley et al. 2006; Ma et al. 2006), a physically based slip evolution is influenced by higher-order stress contribution which is written in terms of gradients of plastic slip, where the arguments of virtual work have not been applied. These extended crystal plasticity models could further be classified according to chosen global degrees of freedom for the finite element solution procedure. In this context, there have been two different approaches where authors take plastic slip (see, e.g., Borg 2007; Yalcinkaya 2011; Yalcinkaya et al. 2012; Klusemann and Yalçinkaya 2013; Klusemann et al. 2013) or the dislocation density (see, e.g., Evers et al. 2004; Bayley et al. 2006; Geers et al. 2007) as degrees of freedom together with displacement field. For the sake of simplicity, the current study considers the most elementary representation of a strain gradient crystal plasticity theory in terms of plastic slip and displacement which are taken as coupled global degrees of freedom.

In addition to their success in predicting size effects, strain gradient crystal plasticity models have been improved further to simulate the intragranular microstructure evolution, intergranular grain boundary behavior, and macroscopic strain localization leading to necking in metallic materials (see, e.g., also Yalcinkaya et al. 2011, 2012; Özdemir and Yalçinkaya 2014; Yalcinkaya and Lancioni 2014; Lancioni et al. 2015a). A thermodynamically consistent incorporation of a proper plastic potentials results in different responses in terms of both strain distribution and the global constitutive response. This chapter addresses the thermodynamically consistent development and extension of higher-order, work-conjugate, plastic slip-based strain gradient crystal plasticity frameworks to capture the size effect and inhomogeneous strain and stress distribution in polycrystalline materials, in a simple and illustrative setting. It also studies different approaches to develop these models, the boundary conditions, and the solution algorithms and addresses some future challenges.

The chapter is organized as follows. First, in section “[Plastic Slip-Based Strain Gradient Crystal Plasticity](#),” the classical thermodynamical derivation of the rate-dependent strain gradient crystal plasticity framework is discussed. Then in section “[Rate Variational Formulation of Strain Gradient Crystal Plasticity](#),” the rate variational formulation of the same model is addressed. In section “[Finite Element Solution Procedure of Strain Gradient Crystal Plasticity Framework](#),” the finite element solution procedure of the framework is studied in detail. Then in section “[Simulation of Polycrystalline Behavior](#),” numerical examples are solved to illustrate the size effect in polycrystalline metals. Last, in section “[Conclusion and Outlook](#)” the concluding remarks are summarized.

Plastic Slip-Based Strain Gradient Crystal Plasticity

In this section the complete formulation and thermodynamical consistency of the higher-order, work-conjugate, plastic slip-based rate-dependent strain gradient crystal plasticity model are studied, and the weak form of equations is derived for the finite element implementation. Depending on the plastic slip potential

for the energetic hardening, the models are distinguished as convex and non-convex in nature. The convex type of model, which is used in this chapter, is essentially used for the size effect predictions, while the non-convex type of models could be employed for the simulation of intrinsic microstructure evolution or macroscopic localization and necking (see, e.g., Yalcinkaya et al. 2011, 2012; Yalçinkaya 2013; Yalcinkaya and Lancioni 2014). The purpose of this section is to study the thermodynamics and the derivation of the convex-type strain gradient crystal plasticity model and to address the ways to incorporate different physical phenomena into the developed framework.

The theoretical framework is developed in a geometrically linear context, with small displacements, strains, and rotations. The time-dependent displacement field is denoted by $\mathbf{u} = \mathbf{u}(\mathbf{x}, t)$, where the vector \mathbf{x} indicates the position of a material point. The strain tensor $\boldsymbol{\varepsilon}$ is defined as $\boldsymbol{\varepsilon} = \frac{1}{2}(\nabla\mathbf{u} + (\nabla\mathbf{u})^T)$, and the velocity vector is represented as $\mathbf{v} = \dot{\mathbf{u}}$. The strain is decomposed additively as

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}^e + \boldsymbol{\varepsilon}^p \quad (1)$$

into an elastic part $\boldsymbol{\varepsilon}^e$ and a plastic part $\boldsymbol{\varepsilon}^p$. The plastic strain rate can be written as a summation of plastic slip rates on the individual slip systems, $\dot{\boldsymbol{\varepsilon}}^p = \sum_{\alpha} \dot{\gamma}^{\alpha} \mathbf{P}^{\alpha}$ with $\mathbf{P}^{\alpha} = \frac{1}{2}(\mathbf{s}^{\alpha} \otimes \mathbf{n}^{\alpha} + \mathbf{n}^{\alpha} \otimes \mathbf{s}^{\alpha})$ the symmetrized Schmid tensor, where \mathbf{s}^{α} and \mathbf{n}^{α} are the unit slip direction vector and unit normal vector on slip system α , respectively.

Next, the selection of the internal state variables is discussed. The choice of the state variables controls the formulation of the free energy and the governing equations which would be needed to obtain the constitutive response and the evolution of the deformation field. The state variables are chosen to be given by the set

$$state = \{\boldsymbol{\varepsilon}^e, \gamma^{\alpha}, \nabla\gamma^{\alpha}\} \quad (2)$$

where γ^{α} contains the plastic slips on the different slip systems α and $\nabla\gamma^{\alpha}$ represents the gradient of the slips on these slip systems. The chosen variables describing the state are to be regarded as mesoscale internal state variables. At that scale, the glide plane slip and their gradients are natural candidates. They both describe the physical state on a slip plane in the mean field sense. The slip characterizes the average plastic deformation accumulated on a glide plane, whereas their gradients characterize the amount of geometrically necessary dislocations that accompany that process, also an important characteristic of the mean dislocation configuration. The use of these state variables naturally entails the other quantities in the constitutive description, like the gradient of the dislocation density, where the divergence of the microstress directly involves the gradient of GND.

There are a number of choices possible for the state variables. At the mesoscale, the choice made is quite appropriate in representing a spatiotemporal ensemble of microscopic states. The mesoscale state variables are measures of the current dislocation state in the material relative to which energy storage and hardening in the

material are modeled. Microscopic processes like phase transitions or dislocation interaction may survive coarse-graining (from the microscale to the mesoscale), resulting in non-convex contributions to the free energy as in phase field models and a transition from spatial or material homogeneity or inhomogeneity. The use of slip-like quantities as mesoscale state variables is quite common in crystal plasticity. Several examples can be found in the literature, e.g., Rice (1971), using a continuum slip model, which characterizes the state of the crystal in terms of the shear strains on each slip system. Later on Perzyna (1988) considered plastic slips together with the slip resistance as internal state variables. For the numerical problems with monotonic loading histories, the adopted mesoscopic state variables are well capable of bridging the microscopic and mesoscopic states of the material.

Following the arguments of Gurtin (e.g., Gurtin 2000, 2002), the power expended by each independent rate-like kinematical descriptor is expressible in terms of an associated force consistent with its own balance. However, the basic kinematical fields of rate variables, namely, $\dot{\mathbf{e}}^e$, $\dot{\mathbf{u}}$, and $\dot{\gamma}^\alpha$ are not spatially independent. It is therefore not immediately clear how the associated force balances are to be formulated, and, for that reason, these balances are established using the principle of virtual power.

Assuming that at a fixed time the fields \mathbf{u} , \mathbf{e}^e , and γ^α are known, we consider $\delta\dot{\mathbf{u}}$, $\delta\dot{\mathbf{e}}^e$, and $\delta\dot{\gamma}^\alpha$ as virtual rates, which are collected in the generalized virtual velocity $\mathcal{V} = \{\delta\dot{\mathbf{u}}, \delta\dot{\mathbf{e}}^e, \delta\dot{\gamma}^\alpha\}$. \mathcal{P}_{ext} is the power expended on the domain Ω and \mathcal{P}_{int} a concomitant expenditure of power within Ω , given by

$$\begin{aligned}\mathcal{P}_{ext}(\Omega, \mathcal{V}) &= \int_S \mathbf{t}(\mathbf{n}) \cdot \delta\dot{\mathbf{u}} \, dS + \int_S \sum_\alpha \chi^\alpha(\mathbf{n}) \delta\dot{\gamma}^\alpha \, dS \\ \mathcal{P}_{int}(\Omega, \mathcal{V}) &= \int_\Omega \boldsymbol{\sigma} : \delta\dot{\mathbf{e}}^e \, d\Omega + \int_\Omega \sum_\alpha (\pi^\alpha \delta\dot{\gamma}^\alpha) \, d\Omega + \int_\Omega \sum_\alpha (\boldsymbol{\xi}^\alpha \cdot \nabla \delta\dot{\gamma}^\alpha) \, d\Omega\end{aligned}\quad (3)$$

where the stress tensor $\boldsymbol{\sigma}$, the scalar internal forces π^α , and the microstress vectors $\boldsymbol{\xi}^\alpha$ are the thermodynamical forces conjugate to the internal state variables \mathbf{e}^e , γ^α , and $\nabla\gamma^\alpha$, respectively. In \mathcal{P}_{ext} , $\mathbf{t}(\mathbf{n})$ is the macroscopic surface traction, while $\chi^\alpha(\mathbf{n})$ represents the microscopic surface traction conjugate to γ^α at the boundary S with \mathbf{n} indicating the boundary normal. The principle of virtual power states that for any generalized virtual velocity \mathcal{V} , the corresponding internal and external power are balanced, i.e.,

$$\mathcal{P}_{ext}(\Omega, \mathcal{V}) = \mathcal{P}_{int}(\Omega, \mathcal{V}) \quad \forall \mathcal{V} \quad (4)$$

Now the consequences are derived following the restriction $\nabla\delta\dot{\mathbf{u}} = \delta\dot{\mathbf{e}}^e + \delta\dot{\boldsymbol{\omega}}^e + \sum_\alpha \delta\dot{\gamma}^\alpha \mathbf{s}^\alpha \otimes \mathbf{m}^\alpha$. First a generalized virtual velocity without slip rate is considered, namely, $\delta\dot{\gamma}^\alpha = 0$ which means $\nabla\delta\dot{\mathbf{u}} = \delta\dot{\mathbf{e}}^e + \delta\dot{\boldsymbol{\omega}}^e$. Considering that Cauchy stress is symmetric, the power balance becomes

$$\int_{\Omega} \boldsymbol{\sigma} : \nabla \delta \dot{\mathbf{u}} d\Omega = \int_S \mathbf{t}(\mathbf{n}) \cdot \delta \dot{\mathbf{u}} dS \quad (5)$$

After applying divergence theorem, the conditions derived from equation (5) are the traction condition

$$\mathbf{t}(\mathbf{n}) = \boldsymbol{\sigma} \mathbf{n} \quad (6)$$

and the classical linear momentum balance

$$\nabla \cdot \boldsymbol{\sigma} = \mathbf{0}. \quad (7)$$

The microscopic counterparts of the conditions are obtained through a consideration of a generalized virtual velocity with $\delta \dot{\mathbf{u}} = \mathbf{0}$ with arbitrary $\delta \dot{\gamma}^\alpha$ field which results in

$$\sum_{\alpha} \delta \dot{\gamma}^\alpha (\mathbf{s}^\alpha \otimes \mathbf{m}^\alpha) = -\delta \dot{\boldsymbol{\varepsilon}}^e - \delta \dot{\boldsymbol{\omega}}^e \quad (8)$$

And the term $\boldsymbol{\sigma} : \delta \dot{\boldsymbol{\varepsilon}}^e$ becomes

$$\boldsymbol{\sigma} : \delta \dot{\boldsymbol{\varepsilon}}^e = -\boldsymbol{\sigma} : \sum_{\alpha} \delta \dot{\gamma}^\alpha (\mathbf{s}^\alpha \otimes \mathbf{m}^\alpha) - \boldsymbol{\sigma} : \delta \dot{\boldsymbol{\omega}}^e \quad (9)$$

Using the symmetry of the Cauchy stress, the following relation is obtained $\boldsymbol{\sigma} : \delta \dot{\boldsymbol{\varepsilon}}^e = -\sum_{\alpha} \delta \dot{\gamma}^\alpha \tau^\alpha$ where τ^α is the Schmid resolved stress. For this case the power balance (4) utilized again

$$\begin{aligned} \int_S \sum_{\alpha} (\chi^\alpha(\mathbf{n}) \delta \dot{\gamma}^\alpha) dS &= -\sum_{\alpha} \int_{\Omega} \tau^\alpha \delta \dot{\gamma}^\alpha d\Omega + \sum_{\alpha} \int_{\Omega} \pi^\alpha \delta \dot{\gamma}^\alpha d\Omega \\ &+ \sum_{\alpha} \int_{\Omega} \boldsymbol{\xi}^\alpha \cdot \nabla \delta \dot{\gamma}^\alpha d\Omega \end{aligned} \quad (10)$$

or equivalently

$$\sum_{\alpha} \int_{\Omega} (-\tau^\alpha + \pi^\alpha - \nabla \cdot \boldsymbol{\xi}^\alpha) \delta \dot{\gamma}^\alpha d\Omega + \sum_{\alpha} \int_S (\boldsymbol{\xi}^\alpha \cdot \mathbf{n} - \chi^\alpha(\mathbf{n})) \delta \dot{\gamma}^\alpha dS = 0 \quad (11)$$

which should be satisfied for all $\delta \dot{\gamma}^\alpha$. This argument yields the microscopic traction condition on the outer boundary of the bulk material

$$\chi^\alpha(\mathbf{n}) = \boldsymbol{\xi}^\alpha \cdot \mathbf{n} \quad (12)$$

and the microscopic force balance inside the bulk material on each slip system α

$$\tau^\alpha - \pi^\alpha + \nabla \cdot \boldsymbol{\xi}^\alpha = 0 \quad (13)$$

on each slip system α .

Next, the thermodynamical consistent derivation of global system of equations is addressed, starting with the local internal power expression

$$P_i = \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}^e + \sum_{\alpha} (\pi^{\alpha} \dot{\gamma}^{\alpha} + \boldsymbol{\xi}^{\alpha} \cdot \nabla \dot{\gamma}^{\alpha}) \quad (14)$$

The local dissipation inequality results in

$$D = P_i - \dot{\psi} = \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}^e + \sum_{\alpha} (\pi^{\alpha} \dot{\gamma}^{\alpha} + \boldsymbol{\xi}^{\alpha} \cdot \nabla \dot{\gamma}^{\alpha}) - \dot{\psi} \geq 0 \quad (15)$$

For a single crystal material without grain boundaries, the material is assumed to be endowed with a free energy with different contributions according to

$$\psi = \psi_e + \psi_{\gamma} + \psi_{\nabla\gamma} \quad (16)$$

The time derivative of the free energy is expanded and equation (15) is elaborated to

$$\begin{aligned} D &= \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}^e + \sum_{\alpha} (\pi^{\alpha} \dot{\gamma}^{\alpha} + \boldsymbol{\xi}^{\alpha} \cdot \nabla \dot{\gamma}^{\alpha}) - \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}^e} : \dot{\boldsymbol{\varepsilon}}^e - \frac{\partial \psi}{\partial \gamma^{\alpha}} \dot{\gamma}^{\alpha} - \frac{\partial \psi}{\partial \nabla \gamma^{\alpha}} \cdot \nabla \dot{\gamma}^{\alpha} \\ &= \underbrace{\left(\boldsymbol{\sigma} - \frac{d\psi_e}{d\boldsymbol{\varepsilon}^e} \right)}_0 : \dot{\boldsymbol{\varepsilon}}^e + \sum_{\alpha} \left(\pi^{\alpha} - \frac{\partial \psi_{\gamma}}{\partial \gamma^{\alpha}} \right) \dot{\gamma}^{\alpha} + \sum_{\alpha} \underbrace{\left(\boldsymbol{\xi}^{\alpha} - \frac{\partial \psi_{\nabla\gamma}}{\partial \nabla \gamma^{\alpha}} \right)}_0 \cdot \nabla \dot{\gamma}^{\alpha} \geq 0 \end{aligned} \quad (17)$$

The stress $\boldsymbol{\sigma}$ and the microstress vectors $\boldsymbol{\xi}^{\alpha}$ are regarded as energetic quantities having no contribution to the dissipation

$$\begin{aligned} \boldsymbol{\sigma} &= \frac{d\psi_e}{d\boldsymbol{\varepsilon}^e} \\ \boldsymbol{\xi}^{\alpha} &= \frac{\partial \psi_{\nabla\gamma}}{\partial \nabla \gamma^{\alpha}} \end{aligned} \quad (18)$$

whereas π^{α} does have a dissipative contribution. Note that considering the microstress vectors, $\boldsymbol{\xi}^{\alpha}$ as non-dissipative is an assumption. It is also possible to take dissipative and non-dissipative parts of the microstresses due to the gradient of the plastic deformation as done in some of the recent works (see, e.g., Gurtin 2002, 2008). In that case the dissipative parts would appear in the following reduced dissipation inequality, while in the current framework it does not exist

$$D = \sum_{\alpha} \left(\pi^{\alpha} - \frac{\partial \psi_{\gamma}}{\partial \gamma^{\alpha}} \right) \dot{\gamma}^{\alpha} \geq 0. \quad (19)$$

The multipliers of the plastic slip rates are identified as the set of dissipative stresses $\sigma_{\text{dis}}^\alpha$

$$\sigma_{\text{dis}}^\alpha = \pi^\alpha - \frac{\partial \psi_\gamma}{\partial \gamma^\alpha} \quad (20)$$

In order to satisfy the reduced dissipation inequality at the slip system level, the following constitutive equation is proposed

$$\sigma_{\text{dis}}^\alpha = \varphi^\alpha \text{sign}(\dot{\gamma}^\alpha) \quad (21)$$

where φ^α represents the mobilized slip resistance of the slip system under consideration

$$\varphi^\alpha = \frac{s^\alpha}{\dot{\gamma}_0^\alpha} |\dot{\gamma}^\alpha| \quad (22)$$

where s^α is the resistance to dislocation slip which is assumed to be constant and $\dot{\gamma}_0$ is the reference slip rate. Substituting (22) into (21) gives

$$\dot{\gamma}^\alpha = \frac{\dot{\gamma}_0^\alpha}{s^\alpha} \sigma_{\text{dis}}^\alpha \quad (23)$$

Substitution of $\sigma_{\text{dis}}^\alpha$ according to (20) into (23) reveals

$$\dot{\gamma}^\alpha = \frac{\dot{\gamma}_0^\alpha}{s^\alpha} \left(\pi^\alpha - \frac{\partial \psi_\gamma}{\partial \gamma^\alpha} \right) \quad (24)$$

Using the microforce balance (13) results in the plastic slip equation

$$\dot{\gamma}^\alpha = \frac{\dot{\gamma}_0^\alpha}{s^\alpha} \left(\tau^\alpha + \nabla \cdot \xi^\alpha - \frac{\partial \psi_\gamma}{\partial \gamma^\alpha} \right). \quad (25)$$

Note that the derivation of the thermodynamically consistent constitutive equations (plastic slip evolution) are based on an assumption which results in an expression identical to classical power-law relation of crystal plasticity frameworks with an exponent $m = 1$. This choice was of course made for simplicity only, yet with a large similarity to discrete dislocation studies using linear drag relations. The general form of equations would follow by including the rate sensitivity exponent in (22) as $\varphi^\alpha = s^\alpha (\dot{\gamma}_0^\alpha / |\dot{\gamma}^\alpha|)^m$ which would yield the following general form

$$\dot{\gamma}^\alpha = \dot{\gamma}_0^\alpha \left(\left| \tau^\alpha + \nabla \cdot \xi^\alpha - \frac{\partial \psi_\gamma}{\partial \gamma^\alpha} \right| / s^\alpha \right)^{\frac{1}{m}} \text{sign} \left(\tau^\alpha - \frac{\partial \psi_\gamma}{\partial \gamma^\alpha} \right) \quad (26)$$

The identification of the model as crystal plasticity, strain gradient crystal plasticity, or non-convex strain gradient crystal plasticity depends on the terms entering the equations (25) and (26). Comparing equations (23) and (25) resolves $\sigma_{\text{dis}}^\alpha = \tau^\alpha + \nabla \cdot \xi^\alpha - \partial\psi_\gamma/\partial\gamma^\alpha$. In the recently developed non-convex strain gradient crystal plasticity frameworks (see, e.g., Yalcinkaya et al. 2011, 2012), the driving force for the dislocation slip evolution is $\sigma_{\text{dis}}^\alpha$ which physically means that, in addition to the resolved shear stress τ^α , the back stress due to the gradients of the geometrically necessary dislocation densities $\nabla \cdot \xi^\alpha$, and the internal force leading to the accumulation of plastic slip $\partial\psi_\gamma/\partial\gamma^\alpha$, is affecting the plastic flow. In classical crystal plasticity frameworks, it is only the resolved shear stress τ^α which determines the plastic flow, while in strain gradient type of models it is the effective resolved shear stress $\tau_{\text{eff}}^\alpha = \tau^\alpha + \nabla \cdot \xi^\alpha$. In addition to the explicit contribution of ψ_γ , other contributions of the free energies defined in (16) enter the slip equation via (18) with $\tau^\alpha = d\psi_e/d\boldsymbol{\varepsilon}^e : \mathbf{P}^\alpha$ and $\xi^\alpha = \partial\psi_{\nabla\gamma}/\partial\nabla\gamma^\alpha$.

It is necessary to comment on the particular choices for energy potentials ($\psi_e, \psi_\gamma, \psi_{\nabla\gamma}$) which would result in the expressions for the related stresses with $\tau^\alpha = d\psi_e/d\boldsymbol{\varepsilon}^e : \mathbf{P}^\alpha$, $\xi^\alpha = \partial\psi_{\nabla\gamma}/\partial\nabla\gamma^\alpha$ and $\partial\psi_\gamma/\partial\gamma^\alpha$. In the numerical examples of this chapter, quadratic forms are used for the elastic free energy ψ_e and the plastic slip gradients free energy contribution $\psi_{\nabla\gamma}$, i.e.,

$$\psi_e = \frac{1}{2} \boldsymbol{\varepsilon}^e : {}^4\mathbf{C} : \boldsymbol{\varepsilon}^e \quad \text{and} \quad \psi_{\nabla\gamma} = \sum_{\alpha} \frac{1}{2} A \nabla\gamma^\alpha \cdot \nabla\gamma^\alpha \quad (27)$$

where A is a scalar quantity, which includes an internal length scale parameter, governing the effect of the plastic slip gradients on the internal stress field, and it could be introduced in different ways. One choice is $A = ER^2/(16(1-\nu^2))$ as, e.g., used in Bayley et al. (2006) and Geers et al. (2007), where R is a typical length scale for dislocation interactions and it physically represents the radius of the dislocation domain contributing to the internal stress field. If the dislocation interaction is limited to nearest neighbor interactions only, then R equals the dislocation spacing. Moreover, ν is Poisson's ratio and E is Young's modulus. Depending on the problem, different relations for A could be introduced or it could even be used as a parameter itself in a more phenomenological way. The definitions in (27) result in the following stress expressions

$$\begin{aligned} \boldsymbol{\sigma} &= \frac{d\psi_e}{d\boldsymbol{\varepsilon}^e} = {}^4\mathbf{C} : \boldsymbol{\varepsilon}^e \\ \xi^\alpha &= \frac{\partial\psi_{\nabla\gamma}}{\partial\nabla\gamma^\alpha} = A \nabla\gamma^\alpha \end{aligned} \quad (28)$$

The contributions from the elastic and gradient potentials constitute the *convex strain gradient crystal plasticity framework*. The crucial part is the determination of the plastic potential ψ_γ which governs the energetic hardening behavior of the model, the spatial distribution, and the localization of the strain and deformation.

While a convex ψ_γ influences solely the hardening behavior of the model, a non-convex contribution incudes deformation localization and macroscopic softening. Due to the incorporation of a non-convex potential, we refer the resulting framework the *non-convex strain gradient crystal plasticity* having the capability to model heterogeneous distribution of deformation (microstructure evolution). The effect of incorporation of both phenomenological and physically based non-convex plastic potentials on the spatial distribution of deformation and the macroscopic constitutive response within the scope of plasticity, damage, and fracture has been illustrated recently in, e.g., Yalcinkaya et al. (2011, 2012), Klusemann and Yalçinkaya (2013), Klusemann et al. (2013), Lancioni et al. (2015a,b).

Another thermodynamically consistent approach to reach the identical constitutive model and plastic slip evolution relation is to work in the framework of continuum thermodynamics and rate variational methods (see, e.g., Svendsen 2004; Reddy 2011a,b) for history-dependent material behavior, which is addressed in the next section. In this case the evolution equations are derived from a dissipation potential. Same results could be obtained through variational formulation and incremental minimization procedure as well (see Lancioni et al. 2015a).

In order to solve initial boundary value problems using the rate-dependent strain gradient crystal plasticity framework, a fully coupled finite element solution algorithm is employed where both the displacement \mathbf{u} and the plastic slips γ^α are considered as primary variables. These fields are determined in the solution domain by solving simultaneously the linear momentum balance (7) and the slip evolution equation (26) for $m = 1$ and constant slip resistance, which constitute the local strong form of the balance equations

$$\begin{aligned} \nabla \cdot \boldsymbol{\sigma} &= \mathbf{0} \\ \dot{\gamma}^\alpha - \frac{\dot{\gamma}_0^\alpha}{s^\alpha} \tau^\alpha - \frac{\dot{\gamma}_0^\alpha}{s^\alpha} \nabla \cdot \boldsymbol{\xi}^\alpha + \frac{\dot{\gamma}_0^\alpha}{s^\alpha} \frac{\partial \psi_\gamma}{\partial \gamma^\alpha} &= 0 \end{aligned} \quad (29)$$

In order to obtain variational expressions representing the weak forms of the governing equations given above, these equations are multiplied by weighting functions $\delta_{\mathbf{u}}$ and δ_γ^α and integrated over the domain Ω . Using the Gauss theorem (S is the boundary of Ω) results in

$$\begin{aligned} G_{\mathbf{u}} &= \int_{\Omega} \nabla \delta_{\mathbf{u}} : \boldsymbol{\sigma} d\Omega - \int_S \delta_{\mathbf{u}} \cdot \mathbf{t} dS \\ G_\gamma^\alpha &= \int_{\Omega} \delta_\gamma^\alpha \dot{\gamma}^\alpha d\Omega - \int_{\Omega} \delta_\gamma^\alpha \frac{\dot{\gamma}_0^\alpha}{s^\alpha} \tau^\alpha d\Omega + \int_{\Omega} \nabla \delta_\gamma^\alpha \cdot \frac{\dot{\gamma}_0^\alpha}{s^\alpha} \boldsymbol{\xi}^\alpha d\Omega \\ &\quad + \int_{\Omega} \delta_\gamma^\alpha \frac{\dot{\gamma}_0^\alpha}{s^\alpha} \frac{\partial \psi_\gamma}{\partial \gamma^\alpha} d\Omega - \int_S \delta_\gamma^\alpha \frac{\dot{\gamma}_0^\alpha}{s^\alpha} \chi^\alpha dS \end{aligned} \quad (30)$$

where \mathbf{t} is the external traction vector on the boundary S , and $\chi^\alpha = \boldsymbol{\xi} \cdot \mathbf{n}$. The domain Ω is subdivided into finite elements, where the unknown fields of the

displacement and slips and the associated weighting functions within each element are approximated by their nodal values multiplied with the interpolation shape functions stored in the \underline{N}^u and \underline{N}^γ matrices,

$$\begin{aligned} \delta_{\mathbf{u}} &= \underline{N}^u \delta_{\mathbf{u}} & \mathbf{u} &= \underline{N}^u \underline{\mathbf{u}} \\ \delta_{\gamma}^{\alpha} &= \underline{N}^{\gamma} \delta_{\gamma}^{\alpha} & \gamma^{\alpha} &= \underline{N}^{\gamma} \underline{\gamma}^{\alpha} \end{aligned} \tag{31}$$

where $\underline{\mathbf{u}}$, $\delta_{\mathbf{u}}$, $\underline{\gamma}^{\alpha}$, and δ_{γ}^{α} are columns containing the nodal variables. Bilinear interpolation functions for the slip field and quadratic interpolation functions for the displacement field are used. An implicit backward Euler time integration scheme is used for $\dot{\gamma}^{\alpha}$ in a typical time increment $[t_n, t_{n+1}]$ which gives $\dot{\gamma}^{\alpha} = [\gamma_{n+1}^{\alpha} - \gamma_n^{\alpha}] / \Delta t$. The discretized element weak forms read

$$\begin{aligned} G_{\mathbf{u}}^e &= \delta_{\mathbf{u}}^T \left[\int_{\Omega^e} \underline{B}^u \underline{\boldsymbol{\sigma}} d\Omega^e - \int_{S^e} \underline{N}^u \underline{\boldsymbol{\tau}} dS^e \right] \\ G_{\gamma}^{\alpha e} &= \delta_{\gamma}^{\alpha} \left[\int_{\Omega^e} \underline{N}^{\gamma T} \underline{N}^{\gamma} \left[\frac{\gamma_{n+1}^{\alpha} - \gamma_n^{\alpha}}{\Delta t} \right] d\Omega^e - \int_{\Omega^e} \frac{\dot{\gamma}_0^{\alpha}}{s^{\alpha}} \underline{N}^{\gamma T} \underline{\boldsymbol{\tau}}^{\alpha} d\Omega^e \right] \\ &+ \delta_{\gamma}^{\alpha} \frac{\dot{\gamma}_0^{\alpha}}{s^{\alpha}} \left[\int_{\Omega^e} \underline{B}^{\gamma} \underline{\boldsymbol{\xi}}^{\alpha} d\Omega^e + \int_{\Omega^e} \underline{N}^{\gamma T} \frac{\partial \psi_{\gamma^{\alpha}}}{\partial \gamma^{\alpha}} d\Omega^e - \int_{S^e} \underline{N}^{\gamma T} \underline{\boldsymbol{\chi}}^{\alpha} dS^e \right] \end{aligned} \tag{32}$$

The weak forms of the balance equations (32) are linearized with respect to the variations of the primary variables $\underline{\mathbf{u}}$ and $\underline{\gamma}^{\alpha}$ and solved by means of a Newton-Raphson solution scheme for the increments of the displacement field $\Delta \underline{\mathbf{u}}$ and the plastic slips $\Delta \gamma^{\alpha}$. The procedure results in a system of linear equations which can be written in the following matrix format

$$\begin{bmatrix} \underline{K}^{uu} & \underline{K}^{u\gamma} \\ \underline{K}^{\gamma u} & \underline{K}^{\gamma\gamma} \end{bmatrix} \begin{bmatrix} \Delta \underline{\mathbf{u}} \\ \Delta \underline{\gamma}^{\alpha} \end{bmatrix} = \begin{bmatrix} -\underline{R}^u + \underline{R}_u^{ext} \\ -\underline{R}^{\gamma} + \underline{R}_\gamma^{ext} \end{bmatrix} \tag{33}$$

where \underline{K}^{uu} , $\underline{K}^{u\gamma}$, $\underline{K}^{\gamma u}$, and $\underline{K}^{\gamma\gamma}$ represent the global tangent matrices, while \underline{R}^u and \underline{R}^{γ} are the global residual columns. The contributions \underline{R}_u^{ext} and $\underline{R}_\gamma^{ext}$ originate from the boundary terms.

The global degrees of freedom in this framework are the displacement and the plastic slips, in terms of which the boundary conditions are defined. Without an explicit grain boundary model, in the current setting, there are two types of conditions that could be used at grain boundaries during polycrystal simulations. The first one is the soft boundary condition for the plastic slip which does not restrict the transfer of dislocations to the neighboring grain, and the other one is the hard boundary condition which blocks the transmission of the dislocations and results in the boundary layer in terms of plastic slip. The reality is in

between, and in order to model the proper behavior of transmission, emission, and dissociation of dislocations within the grain boundary, an explicit grain boundary model should be included in the framework (see very recent examples Özdemir and Yalçinkaya 2014; van Beers et al. 2015a,b; Gottschalk et al. 2016; Bayerschen et al. 2016). For a more detailed discussion on the GB modelling through strain gradient crystal plasticity, please refer to the upcoming chapter by Yalcinkaya and Ozdemir.

Rate Variational Formulation of Strain Gradient Crystal Plasticity

In this section the derivation of the constitutive equations of the rate-dependent strain gradient crystal plasticity framework is conducted via rate variational formulation. The same set of coupled equations are eventually obtained to be solved through finite element method.

In here, the formulation is carried out in the framework of continuum thermodynamics and rate variational methods (see, e.g., Svendsen 2004; Svendsen and Bargmann 2010) for history-dependent behavior. The attention is confined to quasi-static, infinitesimal deformation processes. Let B represent sample domain with boundary ∂B . Besides the displacement field u , the principle global unknowns are taken as the plastic slips γ^α on each glide system α . The model behavior is identified with energetic and dissipative processes. The energetic processes are represented by the free energy density ψ . In the case of general non-convex gradient crystal plasticity (Yalcinkaya et al. 2011, 2012), this consists of an elastic ψ_e , non-convex ψ_γ , and the gradient $\psi_{\nabla\gamma}$ parts (see equation 16).

The dissipative/kinetic processes are represented by a simple rate-dependent power-law form of dissipation potential φ , which would result in the same constitutive relation for the evolution of plastic slip field

$$\varphi = \sum_{\alpha} \frac{1}{m+1} s \dot{\gamma}_0 \left| \frac{\dot{\gamma}^{\alpha}}{\dot{\gamma}_0} \right|^{m+1} \quad (34)$$

The dissipation potential φ is nonnegative and convex in $\dot{\gamma}$; therefore it satisfies the dissipation principle Silhavy (1997) sufficiently. This form tacitly assumes zero activation energy or stress for initiation of inelastic deformation. Since the current work is concerned with purely qualitative effects, $m = 1$ is chosen for simplicity, analogous to discrete dislocation studies based on linear drag relations. When γ is modeled as a global variable, the current framework results in a Ginzburg-Landau-/phase-field-like relation for γ when ψ_γ is introduced as non-convex. Other choices for m , including those used in classical crystal plasticity studies, would result in a non-Ginzburg-Landau form. Other values of m influence the strength, but not the qualitative effect of the rate-dependent material behavior and microstructure development (Yalcinkaya et al. 2011).

Next, the finite element solution procedure of the continuum thermodynamics variational formulation of the evolution field relations for the initial-boundary value problem is addressed briefly. The formulation begins with the following rate functional

$$R = \int_B r_v dV + \int_{\partial B} r_s dS \tag{35}$$

which is based on the corresponding volumetric and boundary rate potentials

$$\begin{aligned} r_v &:= \zeta_v + \varphi_v \\ r_s &:= \zeta_s + \varphi_s \end{aligned} \tag{36}$$

The energetic and the dissipative terms are represented by ζ and φ , respectively. The energy storage density ζ could be determined by

$$\zeta_v = \frac{\partial \psi}{\partial \nabla \mathbf{u}} \cdot \nabla \dot{\mathbf{u}} + \sum_{\alpha} \frac{\partial \psi}{\partial \gamma^{\alpha}} \dot{\gamma}^{\alpha} + \sum_{\alpha} \frac{\partial \psi}{\partial \nabla \gamma^{\alpha}} \cdot \nabla \dot{\gamma}^{\alpha} \tag{37}$$

Using the definition of the stresses due to elastic and gradient free energy in (18)

$$\zeta_v = \boldsymbol{\sigma} \cdot \nabla \dot{\mathbf{u}} + \sum_{\alpha} \left(\frac{\partial \psi_{\gamma}}{\partial \gamma^{\alpha}} - \tau^{\alpha} \right) \dot{\gamma}^{\alpha} + \sum_{\alpha} \boldsymbol{\xi}^{\alpha} \cdot \nabla \dot{\gamma}^{\alpha} \tag{38}$$

The surface rate potential r_s consists of energetic ζ_s and kinetic or dissipative φ_s parts which are linear and nonlinear, respectively, in the rates $\dot{\mathbf{u}}$ and $\dot{\gamma}^{\alpha}$ and determines the flux boundary conditions

$$-\mathbf{t} = \frac{\partial r_s}{\partial \dot{\mathbf{u}}} \quad \text{and} \quad -\chi^{\alpha} = \frac{\partial r_s}{\partial \dot{\gamma}^{\alpha}} \tag{39}$$

associated with $\dot{\mathbf{u}}$ and $\dot{\gamma}_{\alpha}$, respectively. Considering flux free boundary conditions, the rate potentials are

$$\begin{aligned} r_v &= \boldsymbol{\sigma} \cdot \nabla \dot{\mathbf{u}} + \sum_{\alpha} \left(\left(\frac{\partial \psi_{\gamma}}{\partial \gamma^{\alpha}} - \tau^{\alpha} \right) \dot{\gamma}^{\alpha} + \boldsymbol{\xi}^{\alpha} \cdot \nabla \dot{\gamma}^{\alpha} + \frac{1}{2} s \dot{\gamma}_0 \left| \frac{\dot{\gamma}^{\alpha}}{\dot{\gamma}_0} \right|^2 \right) \\ r_s &= \zeta_s + \varphi_s = 0 \quad (\text{flux - free}). \end{aligned} \tag{40}$$

Now we get the first variation of R in $\dot{\mathbf{u}}$ and $\dot{\gamma}_{\alpha}$ by using integration by parts, the divergence theorem, and the variational derivative $\frac{\delta f}{\delta x} = \frac{\partial f}{\partial x} - \nabla \cdot \frac{\partial f}{\partial \nabla x}$

$$\begin{aligned} \delta R = & \int_B \frac{\delta r_v}{\delta \dot{\mathbf{u}}} \delta \dot{\mathbf{u}} dV + \int_{\partial B} \left(\frac{\partial r_v}{\partial \nabla \dot{\mathbf{u}}} \mathbf{n} + \frac{\partial r_s}{\partial \dot{\mathbf{u}}} \right) \cdot \delta \dot{\mathbf{u}} dS \\ & + \sum_{\alpha} \left\{ \int_B \frac{\delta r_v}{\delta \dot{\gamma}^{\alpha}} \delta \dot{\gamma}^{\alpha} dV + \int_{\partial B} \left(\frac{\partial r_v}{\partial \nabla \dot{\gamma}^{\alpha}} \cdot \mathbf{n} + \frac{\partial r_s}{\partial \dot{\gamma}^{\alpha}} \right) \delta \dot{\gamma}^{\alpha} dS \right\} \end{aligned} \quad (41)$$

Note that this form is independent of the gradients of $\delta \dot{\mathbf{u}}$ and $\delta \dot{\gamma}_{\alpha}$.

R is stationary with respect to all admissible variations of $\dot{\mathbf{u}}$ when

$$\begin{aligned} \frac{\delta r_v}{\delta \dot{\mathbf{u}}} &= \mathbf{0} & \text{in } B \\ \frac{\partial r_v}{\partial \nabla \dot{\mathbf{u}}} \mathbf{n} + \frac{\partial r_s}{\partial \dot{\mathbf{u}}} &= \mathbf{0} & \text{on } \partial B_t \end{aligned} \quad (42)$$

hold in the bulk and at the flux part of the boundary ∂B_t . These relations represent the momentum balance in B and on ∂B_t in (rate) variational form.

The stationarity of R with respect to admissible variations of $\dot{\gamma}_{\alpha}$ is given when

$$\begin{aligned} \frac{\delta r_v}{\delta \dot{\gamma}^{\alpha}} &= 0 & \text{in } B \\ \frac{\partial r_v}{\partial \nabla \dot{\gamma}^{\alpha}} \cdot \mathbf{n} + \frac{\partial r_s}{\partial \dot{\gamma}^{\alpha}} &= 0 & \text{on } \partial B_{\varphi} \end{aligned} \quad (43)$$

hold. The physical interpretation of these relations would simply be the generalized flow rule on each glide system.

Finally the field relations for the deformation

$$\begin{aligned} \nabla \cdot \boldsymbol{\sigma} &= \mathbf{0} & \text{in } B \\ \boldsymbol{\sigma} \mathbf{n} &= \mathbf{t} & \text{on } \partial B_t \end{aligned} \quad (44)$$

and for each glide system

$$\begin{aligned} \frac{s}{\dot{\gamma}_0^{\alpha}} \dot{\gamma}^{\alpha} - \tau^{\alpha} - \nabla \cdot \boldsymbol{\xi}^{\alpha} + \frac{\partial \psi_{\gamma}}{\partial \dot{\gamma}^{\alpha}} &= \mathbf{0} & \text{in } B \\ \boldsymbol{\xi}^{\alpha} \cdot \mathbf{n} &= \chi^{\alpha} & \text{on } \partial B_{\varphi} \end{aligned} \quad (45)$$

are obtained. Equations (44) and (45) form the same set of strong form of equations of the previous section (29) to be solved by the finite element method. We follow the same strategy to obtain the weak form of the equations and to solve numerically for the macroscopic response and the evolution of the deformation field. Both approaches result in a thermodynamically consistent strain gradient crystal plasticity model. For the sake of mechanistic understanding, the constitutive relation/the dissipation potential is chosen to be the most simple one; however other choices

could easily be implemented in the framework as well following the steps above. After studying the thermodynamics and formulation of single crystal strain gradient plasticity, we focus on the details of the coupled finite element implementation of the model in the next section.

Finite Element Solution Procedure of Strain Gradient Crystal Plasticity Framework

In this section the detailed finite element implementation of the coupled slip-based rate-dependent (convex) strain gradient crystal plasticity model with three slip systems is presented. Considering constant slip resistance, without slip interactions, and the linear case with $m = 1$, the plastic slip-dependent part of the free energy ψ_γ drops, and slip equation becomes

$$\underline{\dot{\gamma}} - \frac{\dot{\gamma}_0}{s} \underline{\tau} - \frac{\dot{\gamma}_0}{s} \nabla \cdot \underline{\xi} = 0 \quad (46)$$

Substituting $\underline{\xi}$ and collecting constants in matrix form results in

$$\underline{\dot{\gamma}} - \underline{C} \underline{\tau} - \nabla \cdot (\underline{D} \nabla \underline{\gamma}) = 0 \quad (47)$$

where the plastic slips, resolved Schmid stress, and the gradient of the slip systems are included as

$$\underline{\gamma} = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix} \quad \underline{\tau} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \end{bmatrix} \quad \nabla \underline{\gamma} = \begin{bmatrix} \frac{\partial \gamma_1}{\partial x} & \frac{\partial \gamma_1}{\partial y} \\ \frac{\partial \gamma_2}{\partial x} & \frac{\partial \gamma_2}{\partial y} \\ \frac{\partial \gamma_3}{\partial x} & \frac{\partial \gamma_3}{\partial y} \end{bmatrix} \quad (48)$$

and the parameters read

$$\underline{C} = \begin{bmatrix} \frac{\dot{\gamma}_0}{s} & 0 & 0 \\ 0 & \frac{\dot{\gamma}_0}{s} & 0 \\ 0 & 0 & \frac{\dot{\gamma}_0}{s} \end{bmatrix} \quad \underline{D} = \begin{bmatrix} \frac{\dot{\gamma}_0}{s} A & 0 & 0 \\ 0 & \frac{\dot{\gamma}_0}{s} A & 0 \\ 0 & 0 & \frac{\dot{\gamma}_0}{s} A \end{bmatrix} \quad (49)$$

Together with the slip equation in (47), the linear momentum balance

$$\nabla \cdot \underline{\sigma} = \mathbf{0} \quad (50)$$

enters the finite element solution procedure as strong form equations. The weak forms of the equations are obtained in a standard manner, using a Galerkin procedure. Firstly, the balance of linear momentum is tested with a field of virtual displacements δ_u and integrated over the domain Ω , which results in G_u

$$G_u = \int_{\Omega} \delta_u \cdot (\nabla \cdot \boldsymbol{\sigma}) d\Omega \quad (51)$$

which could be written as

$$G_u = \int_{\Omega} \nabla \delta_u : \boldsymbol{\sigma} d\Omega - \int_{\Gamma} \boldsymbol{t} \cdot \delta_u d\Gamma \quad (52)$$

where \boldsymbol{t} is the traction vector $\boldsymbol{t} = \boldsymbol{\sigma} \cdot \boldsymbol{n}$. Then, G_γ is obtained in a similar way by testing the slip equation with virtual plastic slip and integrating over the domain

$$G_\gamma = \int_{\Omega} \dot{\underline{\gamma}} \cdot \underline{\delta}_\gamma d\Omega - \int_{\Omega} \underline{\mathbf{C}} \underline{\boldsymbol{\tau}} \cdot \underline{\delta}_\gamma d\Omega - \int_{\Omega} [\nabla \cdot (\underline{\mathbf{D}} \nabla \underline{\gamma})] \cdot \underline{\delta}_\gamma d\Omega \quad (53)$$

which could be written as

$$G_\gamma = \int_{\Omega} \dot{\underline{\gamma}} \cdot \underline{\delta}_\gamma d\Omega - \int_{\Omega} \underline{\mathbf{C}} \underline{\boldsymbol{\tau}} \cdot \underline{\delta}_\gamma d\Omega + \int_{\Omega} \nabla \underline{\delta}_\gamma : \underline{\mathbf{D}} \nabla \underline{\gamma} d\Omega - \int_{\Gamma} \underline{\boldsymbol{t}} \cdot \underline{\delta}_\gamma d\Gamma \quad (54)$$

with

$$\underline{\boldsymbol{t}} = \underline{\mathbf{D}} \nabla \underline{\boldsymbol{\gamma}} \cdot \boldsymbol{n} \quad (55)$$

Next, the equations are linearized through Newton-Raphson iterations using the linearization operator Δ which is basically defined as $\Delta_x F = \frac{\partial F}{\partial x} \Delta x$

$$\begin{aligned} \text{Lin} G_u &= \Delta_u G_u + \Delta_\gamma G_u + G_u^* = 0 \\ \text{Lin} G_\gamma &= \Delta_\gamma G_\gamma + \Delta_u G_\gamma + G_\gamma^* = 0 \end{aligned} \quad (56)$$

Now each term is linearized

$$\Delta_u G_u = \int_{\Omega} \frac{\partial \boldsymbol{\sigma}}{\partial \boldsymbol{\varepsilon}} : \Delta_u \boldsymbol{\varepsilon} : \nabla \delta_u d\Omega \quad (57)$$

with $\sigma = {}^4\mathbf{C} : (\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p)$

$$\frac{\partial \sigma}{\partial \boldsymbol{\varepsilon}} = {}^4\mathbf{C} \quad (58)$$

$$\Delta_u \boldsymbol{\varepsilon} = \frac{1}{2}(\nabla \Delta \mathbf{u} + (\nabla \Delta \mathbf{u})^T)$$

and the first term of the linearized displacement Galerkin reads

$$\Delta_\gamma G_u = \int_{\Omega} \nabla \delta_u : {}^4\mathbf{C} : \frac{1}{2}(\nabla \Delta \mathbf{u} + (\nabla \Delta \mathbf{u})^T) d\Omega. \quad (59)$$

Then the second term of the linearized displacement Galerkin is derived as follows

$$\Delta_\gamma G_u = \int_{\Omega} \frac{\partial \sigma}{\partial \gamma} \Delta_\gamma : \nabla \delta_u d\Omega \quad (60)$$

In order to calculate $\frac{\partial \sigma}{\partial \gamma}$ the rate of plastic slip is written as

$$\dot{\boldsymbol{\varepsilon}}^p = \sum_{\alpha} \dot{\gamma}^{\alpha} \mathbf{P}^{\alpha} \quad \text{with} \quad \mathbf{P}^{\alpha} = \frac{1}{2}(\mathbf{s}^{\alpha} \otimes \mathbf{n}^{\alpha} + \mathbf{n}^{\alpha} \otimes \mathbf{s}^{\alpha}) \quad (61)$$

and the integration of the plastic slip rate gives

$$\boldsymbol{\varepsilon}_{n+1}^p = \boldsymbol{\varepsilon}_n^p + \sum_{\alpha} (\gamma_{n+1}^{\alpha} - \gamma_n^{\alpha}) \mathbf{P}^{\alpha} \quad (62)$$

then the regarding derivatives read as

$$\begin{aligned} \frac{\partial \boldsymbol{\varepsilon}^p}{\partial \gamma^{\beta}} &= \mathbf{P}^{\beta} \\ \frac{\partial \sigma}{\partial \gamma^{\beta}} &= \frac{\partial \sigma}{\partial \boldsymbol{\varepsilon}^p} : \frac{\partial \boldsymbol{\varepsilon}^p}{\partial \gamma^{\beta}} = -{}^4\mathbf{C} : \mathbf{P}^{\beta} \end{aligned} \quad (63)$$

Finally the second term of the linearized displacement Galerkin expression is expressed as follows

$$\Delta_\gamma G_u = - \int_{\Omega} \nabla \delta_u : {}^4\mathbf{C} : \mathbf{P}^{\beta} \Delta \gamma^{\beta} d\Omega. \quad (64)$$

Then the terms related to the linearization of $G_{\underline{\gamma}}$ are handled. The operator with respect to the displacement is given by

$$\Delta_u G_{\underline{\gamma}} = - \int_{\Omega} \underline{\underline{C}} \frac{\partial \underline{\underline{\tau}}}{\partial \underline{\underline{\epsilon}}} : \Delta_u \underline{\underline{\epsilon}} \underline{\underline{\delta}}_{\underline{\gamma}} d\Omega \quad (65)$$

By using the derivative $\frac{\partial \underline{\underline{\tau}}^\alpha}{\partial \underline{\underline{\epsilon}}} = {}^4 \underline{\underline{C}} : \underline{\underline{P}}^\alpha$ it could be written as

$$\Delta_u G_{\underline{\gamma}} = - \int_{\Omega} \underline{\underline{C}} \frac{1}{2} (\nabla \Delta_u + (\nabla \Delta_u)^T) : {}^4 \underline{\underline{C}} : \underline{\underline{P}}^\alpha \underline{\underline{\delta}}_{\underline{\gamma}} d\Omega. \quad (66)$$

The last term is

$$\Delta_{\underline{\underline{\gamma}}} G_{\underline{\gamma}} = \int_{\Omega} \Delta_{\underline{\underline{\gamma}}} \dot{\underline{\underline{\gamma}}} \underline{\underline{\delta}}_{\underline{\gamma}} d\Omega - \int_{\Omega} \underline{\underline{C}} \Delta_{\underline{\underline{\tau}}} \underline{\underline{\delta}}_{\underline{\gamma}} d\Omega + \int_{\Omega} \underline{\underline{D}} \Delta_{\underline{\underline{\gamma}}} (\nabla \underline{\underline{\gamma}}) \cdot \nabla \underline{\underline{\delta}}_{\underline{\gamma}} d\Omega \quad (67)$$

The following two terms are necessary

$$\Delta_{\underline{\underline{\tau}}} = \frac{\partial \underline{\underline{\tau}}^\alpha}{\partial \underline{\underline{\gamma}}^\beta} \Delta \underline{\underline{\gamma}} = \frac{\partial}{\partial \underline{\underline{\gamma}}^\beta} (\underline{\underline{\sigma}} : \underline{\underline{P}}^\alpha) \Delta \underline{\underline{\gamma}} = - \underline{\underline{P}}^\alpha : {}^4 \underline{\underline{C}} : \underline{\underline{P}}^\beta \Delta \underline{\underline{\gamma}} \quad (68)$$

$$\Delta_{\underline{\underline{\gamma}}} (\nabla \underline{\underline{\gamma}}) = \nabla \Delta \underline{\underline{\gamma}}$$

Finally, equation (67) is written as

$$\Delta_{\underline{\underline{\gamma}}} G_{\underline{\gamma}} = \int_{\Omega} \frac{1}{\Delta t} \underline{\underline{\delta}}_{\underline{\underline{\gamma}}} \Delta \underline{\underline{\gamma}} d\Omega + \int_{\Omega} \underline{\underline{C}} \underline{\underline{P}}^\alpha : {}^4 \underline{\underline{C}} : \underline{\underline{P}}^\beta \Delta \underline{\underline{\gamma}} \underline{\underline{\delta}}_{\underline{\gamma}} d\Omega + \int_{\Omega} \underline{\underline{D}} \nabla (\Delta \underline{\underline{\gamma}}) \cdot \nabla \underline{\underline{\delta}}_{\underline{\gamma}} d\Omega \quad (69)$$

After the linearization procedure, the above system of equations is discretized. The problem is solved in 2D under plane strain assumption. Quadratic interpolation is used for the displacement field, and linear interpolation is used for the plastic slip field. The degree of freedoms in an element is explicitly illustrated in Fig. 1. The interpolation of the fields is conducted through the matrix of shape functions $\underline{\underline{N}}$

$$\begin{aligned} \underline{\underline{\delta}}_u &= \underline{\underline{N}}^u \underline{\underline{\delta}}_u & \underline{\underline{u}} &= \underline{\underline{N}}^u \underline{\underline{u}} \\ \underline{\underline{\delta}}_{\underline{\underline{\gamma}}} &= \underline{\underline{N}}^\gamma \underline{\underline{\delta}}_{\underline{\underline{\gamma}}} & \underline{\underline{\gamma}} &= \underline{\underline{N}}^\gamma \underline{\underline{\gamma}} \end{aligned} \quad (70)$$

and the interpolations of the gradients of the fields are done through the matrix of shape functions' derivatives

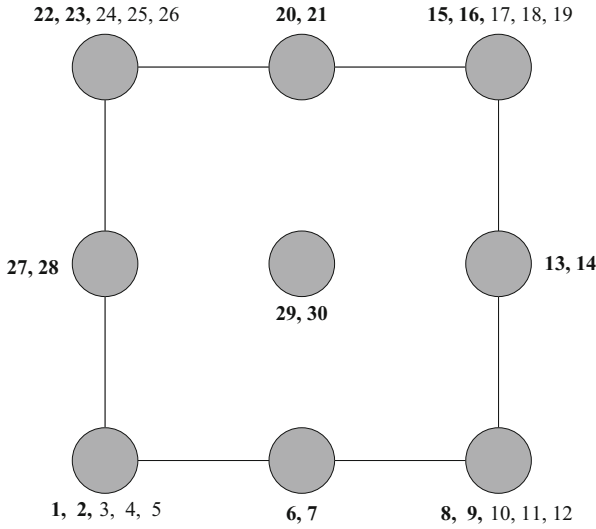


Fig. 1 The displacement degree of freedom numbers is represented with bold, while the other numbers (placed only at the corner nodes) illustrate the plastic dof in an element. This representation corresponds to the quadratic interpolation in \underline{u} and linear interpolation in $\underline{\gamma}$

$$\nabla \delta_{\underline{u}} = \underline{B}^u \delta_{\underline{u}} \quad \nabla \underline{u} = \underline{B}^u \underline{u} \tag{71}$$

$$\nabla \delta_{\underline{\gamma}} = \underline{B}^\gamma \delta_{\underline{\gamma}} \quad \nabla \underline{\gamma} = \underline{B}^\gamma \underline{\gamma}$$

The second-order tensors are stored in Voigt notation as follows: $\underline{A} = [A_{11} \ A_{12} \ A_{21} \ A_{22}]'$. The operations in equations (70) and (71) are explicitly illustrated in the matrix format as follows

$$\begin{bmatrix} u_x \\ u_y \end{bmatrix} = \underbrace{\begin{bmatrix} N_1^u & 0 & N_2^u & 0 & \dots & N_9^u & 0 \\ 0 & N_1^u & 0 & N_2^u & \dots & 0 & N_9^u \end{bmatrix}}_{\underline{N}^u} \begin{bmatrix} u_{1x} \\ u_{1y} \\ u_{2x} \\ u_{2y} \\ \cdot \\ \cdot \\ \cdot \\ u_{9x} \\ u_{9y} \end{bmatrix} \tag{72}$$

$$\begin{bmatrix} \gamma^1 \\ \gamma^2 \\ \gamma^3 \end{bmatrix} = \underbrace{\begin{bmatrix} N_1^\gamma & N_2^\gamma & N_3^\gamma & N_4^\gamma & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & N_1^\gamma & N_2^\gamma & N_3^\gamma & N_4^\gamma & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & N_1^\gamma & N_2^\gamma & N_3^\gamma & N_4^\gamma \end{bmatrix}}_{\underline{N}^\gamma} \begin{bmatrix} \gamma_1^1 \\ \gamma_2^1 \\ \gamma_3^1 \\ \gamma_4^1 \\ \gamma_1^2 \\ \gamma_2^2 \\ \cdot \\ \cdot \\ \gamma_3^3 \\ \gamma_4^3 \end{bmatrix} \quad (73)$$

$$\begin{bmatrix} u_{x,x} \\ u_{x,y} \\ u_{y,x} \\ u_{y,y} \end{bmatrix} = \underbrace{\begin{bmatrix} N_{1,x}^u & 0 & N_{2,x}^u & 0 & \dots & N_{9,x}^u & 0 \\ N_{1,y}^u & 0 & N_{2,y}^u & 0 & \dots & N_{9,y}^u & 0 \\ 0 & N_{1,x}^u & 0 & N_{2,x}^u & \dots & 0 & N_{9,x}^u \\ 0 & N_{1,y}^u & 0 & N_{2,y}^u & \dots & 0 & N_{9,y}^u \end{bmatrix}}_{\underline{B}^u} \begin{bmatrix} u_{1x} \\ u_{1y} \\ u_{2x} \\ u_{2y} \\ \cdot \\ \cdot \\ u_{9x} \\ u_{9y} \end{bmatrix}, \begin{bmatrix} u_{x,x} \\ u_{x,y} \\ u_{y,x} \\ u_{y,y} \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} N_{1,x}^u & 0 & N_{2,x}^u & 0 & \dots & N_{9,x}^u & 0 \\ 0 & N_{1,x}^u & 0 & N_{2,x}^u & \dots & 0 & N_{9,x}^u \\ N_{1,y}^u & 0 & N_{2,y}^u & 0 & \dots & N_{9,y}^u & 0 \\ 0 & N_{1,y}^u & 0 & N_{2,y}^u & \dots & 0 & N_{9,y}^u \end{bmatrix}}_{\underline{B}^{u*}} \begin{bmatrix} u_{1x} \\ u_{1y} \\ u_{2x} \\ u_{2y} \\ \cdot \\ \cdot \\ u_{9x} \\ u_{9y} \end{bmatrix} \quad (74)$$

$$\begin{bmatrix} \gamma_{,1}^1 \\ \gamma_{,2}^1 \\ \gamma_{,1}^2 \\ \gamma_{,2}^2 \\ \gamma_{,1}^3 \\ \gamma_{,2}^3 \end{bmatrix} = \underbrace{\begin{bmatrix} N_{1,x}^\gamma & N_{2,x}^\gamma & N_{3,x}^\gamma & N_{4,x}^\gamma & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ N_{1,y}^\gamma & N_{2,y}^\gamma & N_{3,y}^\gamma & N_{4,y}^\gamma & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & N_{1,x}^\gamma & N_{2,x}^\gamma & N_{3,x}^\gamma & N_{4,x}^\gamma & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & N_{1,y}^\gamma & N_{2,y}^\gamma & N_{3,y}^\gamma & N_{4,y}^\gamma & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & N_{1,x}^\gamma & N_{2,x}^\gamma & N_{3,x}^\gamma & N_{4,x}^\gamma \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & N_{1,y}^\gamma & N_{2,y}^\gamma & N_{3,y}^\gamma & N_{4,y}^\gamma \end{bmatrix}}_{\underline{B}^\gamma} \begin{bmatrix} \gamma_1^1 \\ \gamma_2^1 \\ \gamma_3^1 \\ \gamma_4^1 \\ \gamma_1^2 \\ \gamma_2^2 \\ \cdot \\ \cdot \\ \cdot \\ \gamma_3^3 \\ \gamma_4^3 \end{bmatrix} \tag{75}$$

The following matrix will be needed as well

$$\underline{B}_S^u = \frac{1}{2}(\underline{B}^u + \underline{B}^{u*}) \tag{76}$$

A commonly used operation regarding the fourth- and second-order tensors is $\underline{D} : {}^4C : \underline{B}$ which can be written in index notation as $A = D_{ij} C_{ijkl} B_{kl}$. The following matrix multiplication is used for this operation

$$[A] = \underbrace{\begin{bmatrix} D_{11} & D_{12} & D_{21} & D_{22} \end{bmatrix}}_{\underline{D}^T} \underbrace{\begin{bmatrix} C_{1111} & C_{1112} & C_{1121} & C_{1122} \\ C_{1211} & C_{1212} & C_{1221} & C_{1222} \\ C_{2111} & C_{2112} & C_{2121} & C_{2122} \\ C_{2211} & C_{2212} & C_{2221} & C_{2222} \end{bmatrix}}_{\underline{C}} \underbrace{\begin{bmatrix} B_{11} \\ B_{12} \\ B_{21} \\ B_{22} \end{bmatrix}}_{\underline{B}} \tag{77}$$

After the discretization the linearization operator of the displacement Galerkin could be written as

$$\begin{aligned} \Delta_u G_u &= \int_{\Omega^e} \delta_u^T \underline{B}^{uT} \underline{C} \underline{B}_S^u \Delta u \, d\Omega^e \\ \Delta_\gamma G_u &= - \int_{\Omega^e} \delta_u^T \underline{B}^{uT} \underline{E} \underline{N}^\gamma \Delta \underline{\gamma} \, d\Omega^e \end{aligned} \tag{78}$$

where

$$\begin{aligned}
 \underline{E} &= [\underline{E1} \quad \underline{E2} \quad \underline{E3}] \\
 \underline{E1} &= \underline{C} \underline{P1}^T \\
 \underline{E2} &= \underline{C} \underline{P2}^T \\
 \underline{E3} &= \underline{C} \underline{P3}^T
 \end{aligned} \tag{79}$$

where $\underline{P1}^T$, $\underline{P2}^T$, and $\underline{P3}^T$ represent the column version of the \mathbf{P}^1 , \mathbf{P}^2 , and \mathbf{P}^3 tensors. \underline{C} is the matrix version of the ${}^4\mathbf{C}$ as defined above. The displacement linearization operator of the discretized plastic slip Galerkin reads

$$\Delta_{\mathbf{u}} G_{\gamma} = - \int_{\Omega^e} \delta_{\gamma} \underline{N}^{\gamma T} \underline{C}_l \underline{C}_p \Delta \mathbf{u} d\Omega^e \tag{80}$$

with

$$\underline{C}_l = \begin{bmatrix} \frac{\dot{\gamma}_0}{s} & 0 & 0 \\ 0 & \frac{\dot{\gamma}_0}{s} & 0 \\ 0 & 0 & \frac{\dot{\gamma}_0}{s} \end{bmatrix} \quad \text{and} \quad \underline{C}_p = \begin{bmatrix} \underline{C}_p^1 \\ \underline{C}_p^2 \\ \underline{C}_p^3 \end{bmatrix} \quad \text{and} \quad \begin{aligned} \underline{C}_p^1 &= \underline{P1} \underline{C} \underline{B}_S^u \\ \underline{C}_p^2 &= \underline{P2} \underline{C} \underline{B}_S^u \\ \underline{C}_p^3 &= \underline{P3} \underline{C} \underline{B}_S^u \end{aligned} \tag{81}$$

where \underline{B}_S^u is 4×18 ; \underline{C} is 4×4 ; $\underline{P1}$, $\underline{P2}$, and $\underline{P3}$ are 1×4 ; \underline{C}_p^1 , \underline{C}_p^2 , and \underline{C}_p^3 are 1×18 , \underline{C}_p is 3×18 ; and \underline{C}_l is 3×3 matrices. The slip linearization operator of the discretized plastic slip Galerkin is

$$\begin{aligned}
 \Delta_{\gamma} G_{\gamma} &= \int_{\Omega^e} \frac{1}{\Delta t} \delta_{\gamma} \underline{N}^{\gamma T} \underline{N}^{\gamma} \Delta \underline{\gamma} d\Omega^e + \int_{\Omega^e} \delta_{\gamma} \underline{N}^{\gamma T} \underline{P}^p \underline{N}^{\gamma} \Delta \underline{\gamma} d\Omega^e \\
 &\quad + \int_{\Omega^e} \delta_{\gamma} \underline{B}^{\gamma T} \underline{D}^d \underline{B}^{\gamma} \Delta \underline{\gamma} d\Omega^e
 \end{aligned} \tag{82}$$

where

$$\underline{P}^p = \begin{bmatrix} \frac{\dot{\gamma}_0}{s} \underline{P1} \underline{C} \underline{P1}^T & \frac{\dot{\gamma}_0}{s} \underline{P1} \underline{C} \underline{P2}^T & \frac{\dot{\gamma}_0}{s} \underline{P1} \underline{C} \underline{P3}^T \\ \frac{\dot{\gamma}_0}{s} \underline{P2} \underline{C} \underline{P1}^T & \frac{\dot{\gamma}_0}{s} \underline{P2} \underline{C} \underline{P2}^T & \frac{\dot{\gamma}_0}{s} \underline{P2} \underline{C} \underline{P3}^T \\ \frac{\dot{\gamma}_0}{s} \underline{P3} \underline{C} \underline{P1}^T & \frac{\dot{\gamma}_0}{s} \underline{P3} \underline{C} \underline{P2}^T & \frac{\dot{\gamma}_0}{s} \underline{P3} \underline{C} \underline{P3}^T \end{bmatrix} \quad \text{and} \quad \underline{D}^d$$

$$= \begin{bmatrix} \frac{\dot{\gamma}_0}{s} A & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{\dot{\gamma}_0}{s} A & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\dot{\gamma}_0}{s} A & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\dot{\gamma}_0}{s} A & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{\dot{\gamma}_0}{s} A & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{\dot{\gamma}_0}{s} A & 0 \end{bmatrix} \tag{83}$$

And the discretized Galerkins of previous step are written as

$$\begin{aligned} G_u^* &= \int_{\Omega^e} \delta_u^T \underline{B}_S^u \underline{\sigma} d\Omega^e \\ G_\gamma^* &= \int_{\Omega^e} \delta_\gamma \underline{N}^{\gamma T} \underline{N}^\gamma \left(\frac{\gamma_{n+1} - \gamma_n}{\Delta t} \right) d\Omega^e - \int_{\Omega^e} \delta_\gamma \underline{N}^{\gamma T} \underline{C} \underline{\tau} d\Omega^e \\ &\quad + \int_{\Omega^e} \delta_\gamma \underline{B}^{\gamma T} \underline{D}^d \underline{B}^\gamma \underline{\gamma} d\Omega^e. \end{aligned} \tag{84}$$

Eventually, the element stiffness tangents would be

$$\begin{aligned} k_{uu} &= \int_{\Omega^e} \underline{B}^{uT} \underline{C} \underline{B}_S^u d\Omega^e \\ k_{u\gamma} &= - \int_{\Omega^e} \underline{B}^{uT} \underline{E} \underline{N}^\gamma d\Omega^e \\ k_{\gamma u} &= - \int_{\Omega^e} \underline{N}^{\gamma T} \underline{C}_l \underline{C}_p d\Omega^e \\ k_{\gamma\gamma} &= \int_{\Omega^e} \frac{1}{\Delta t} \underline{N}^{\gamma T} \underline{N}^\gamma \Omega^e + \int_{\Omega^e} \underline{N}^{\gamma T} \underline{P}^p \underline{N}^\gamma d\Omega^e + \int_{\Omega^e} \underline{B}^{\gamma T} \underline{D}^d \underline{B}^\gamma d\Omega^e. \end{aligned} \tag{85}$$

Element residuals are calculated by using the values from the previous estimate

$$\begin{aligned} r_u &= \int_{\Omega^e} \underline{B}_S^u \underline{\sigma} d\Omega^e \\ r_\gamma &= \int_{\Omega^e} \underline{N}^{\gamma T} \underline{N}^\gamma \left(\frac{\gamma_{n+1} - \gamma_n}{\Delta t} \right) d\Omega^e - \int_{\Omega^e} \underline{N}^{\gamma T} \underline{C}_l \underline{\tau} d\Omega^e \\ &\quad + \int_{\Omega^e} \underline{B}^{\gamma T} \underline{D}^d \underline{B}^\gamma \underline{\gamma} d\Omega^e. \end{aligned} \tag{86}$$

The assembly operation gives the global tangent and the residual

$$\begin{aligned} K^{uu} &= \mathbb{A} \{k_{uu}\} & K^{u\gamma} &= \mathbb{A} \{k_{u\gamma}\} & R^u &= \mathbb{A} \{r_u\} \\ K^{\gamma u} &= \mathbb{A} \{k_{\gamma u}\} & K^{\gamma\gamma} &= \mathbb{A} \{k_{\gamma\gamma}\} & R^\gamma &= \mathbb{A} \{r_\gamma\} \end{aligned} \quad (87)$$

and the system of equations to be solved reads

$$\begin{bmatrix} K^{uu} & K^{u\gamma} \\ K^{\gamma u} & K^{\gamma\gamma} \end{bmatrix} \begin{bmatrix} \Delta u \\ \Delta \gamma \end{bmatrix} = \begin{bmatrix} -R^u + R_u^{ext} \\ -R^\gamma + R_\gamma^{ext} \end{bmatrix} \quad (88)$$

where R_u^{ext} and R_γ^{ext} originate from the boundary terms in the equilibrium. Above procedure is implemented in Matlab and Abaqus software, and numerical problems illustrating the behavior of polycrystalline materials are solved and presented in the next section.

Simulation of Polycrystalline Behavior

In this section the performance of the strain gradient crystal plasticity model is tested through size effect simulations using Voronoi tessellation (see, e.g., Aurenhammer 1991), which is used to obtain basic geometries of polycrystalline aggregates containing 14, 110 or 212 grains, presented in Fig. 2. For each aggregate three different average grain sizes are used: $D_{avg} = 50, 100$ and $150 \mu\text{m}$. This compares well with the average grain sizes of AISI 304 stainless steels (1 to $47 \mu\text{m}$, Di Schino and Kenny 2003) and AISI 316L (13 to $107 \mu\text{m}$ Feaugas and Haddou 2003).

Displacement load is applied to the left and right edge in the global -X direction (left edge) and the global +X direction (right edge), resulting in macroscopic $\langle \epsilon_{11} \rangle = 10\%$. The symbol $\langle \rangle$ represents the Macaulay bracket, indicating a macroscopically averaged value. Rigid body movement is prevented by fixing the bottom left and bottom right nodes of the model in global Y direction. If one imagines that the aggregate is embedded in a larger piece of material, the applied boundary condition lies between the two bordering cases of constraints, imposed to the aggregate by the surrounding material: (a) stresses or (b) displacements imposed on the aggregate boundaries. The latter condition results in straight edges. Plastic slip at the boundaries of the model is not constrained.

Material parameters are taken from Yalcinkaya et al. (2012) and are used to demonstrate the strain gradient effects in the polycrystalline aggregates. They are not directly related to any engineering materials. Table 1 lists the used strain gradient crystal plasticity material properties. Crystallographic orientations of grains are randomly distributed ($0-360^\circ$) using uniform probability distribution. Three slip systems are considered in each grain.

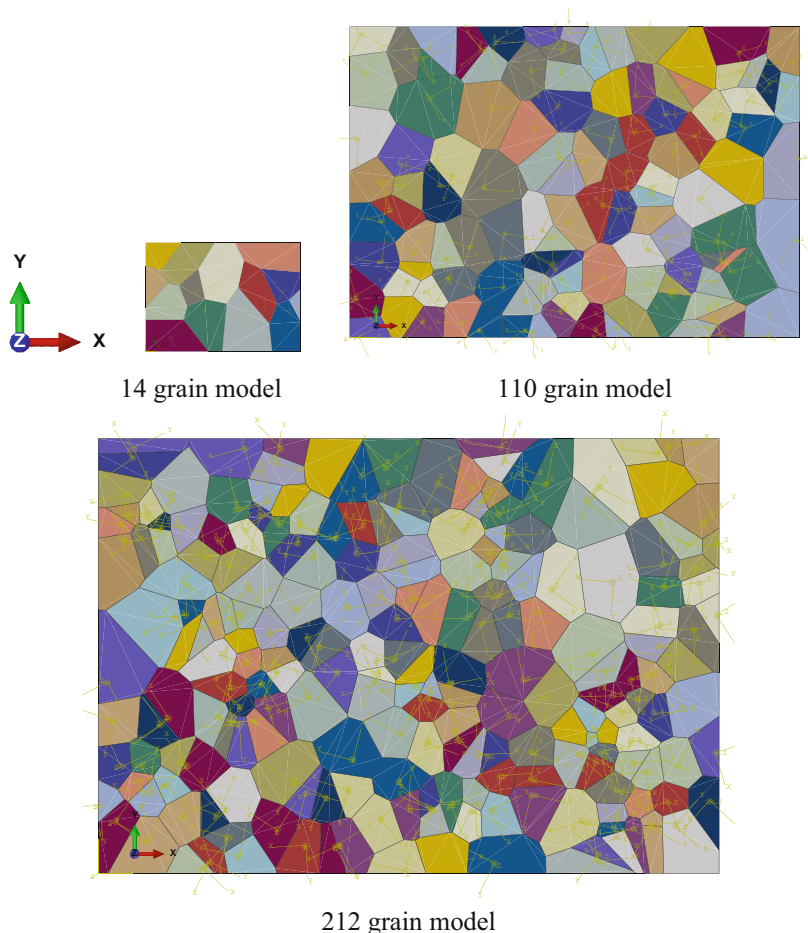


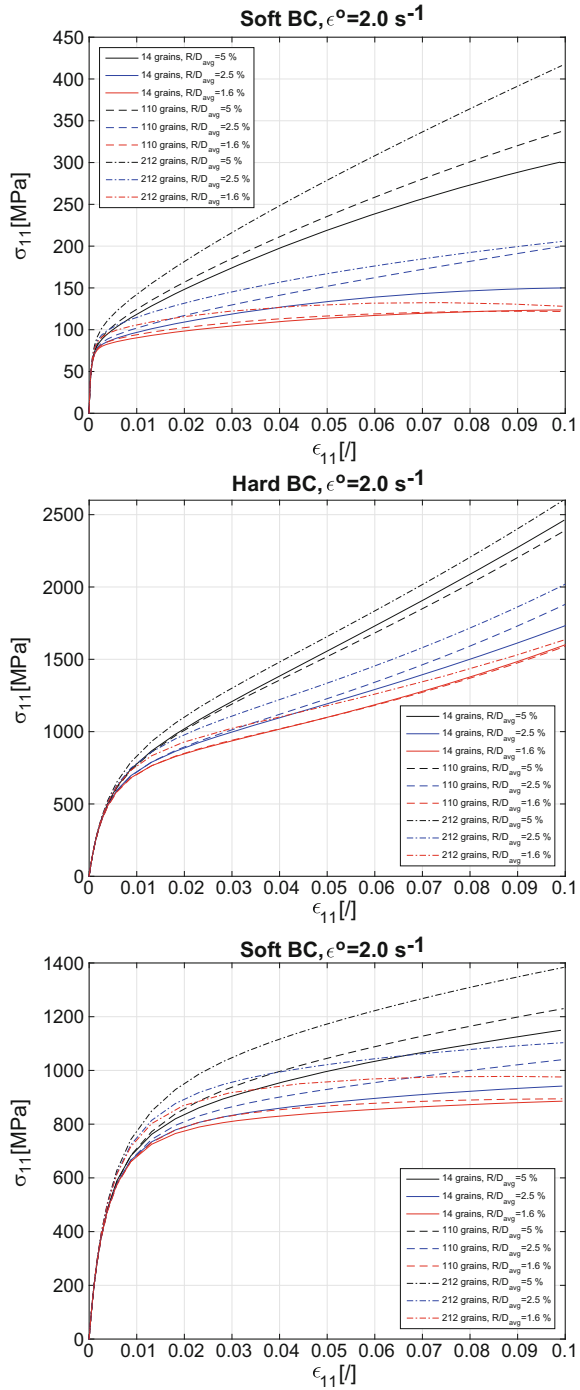
Fig. 2 14, 110, and 212 grain models. Colors represent individual grains; grain orientations are given by local coordinate systems

Table 1 Material properties of the strain gradient single crystal plasticity model

Young modulus	Poisson ratio	Reference slip rate	Slip resistance	Orientations	Material length scale
E [MPa]	ν [/]	$\dot{\gamma}_0$ [s^{-1}]	s [MPa]	[$^\circ$]	R [μm]
210000.0	0.33	0.15	20.0	120, 60, 45	0.0, 1, 5, 7.5, 10

Material length scale parameter R or the average grain size has substantial effect on the plastic behavior of the material. R is related to different microstructural features such as dislocation spacing (see, e.g., Nix and Gao 1998), dislocation source distance (see, e.g., Aifantis et al. 2009), or grain size (see, e.g., Voyiadjis and Abu Al-Rub 2005). Here we relate R to a certain percentage of the average grain size, D_{avg} . The ratio R/D_{avg} has a significant effect on the macroscopic stress-strain response. Larger R/D_{avg} values lead to stiffer responses; see Fig. 3.

Fig. 3 The effect of the material length scale parameter R for strain rates of $\dot{\epsilon} = 0.2$ and 2 s^{-1}



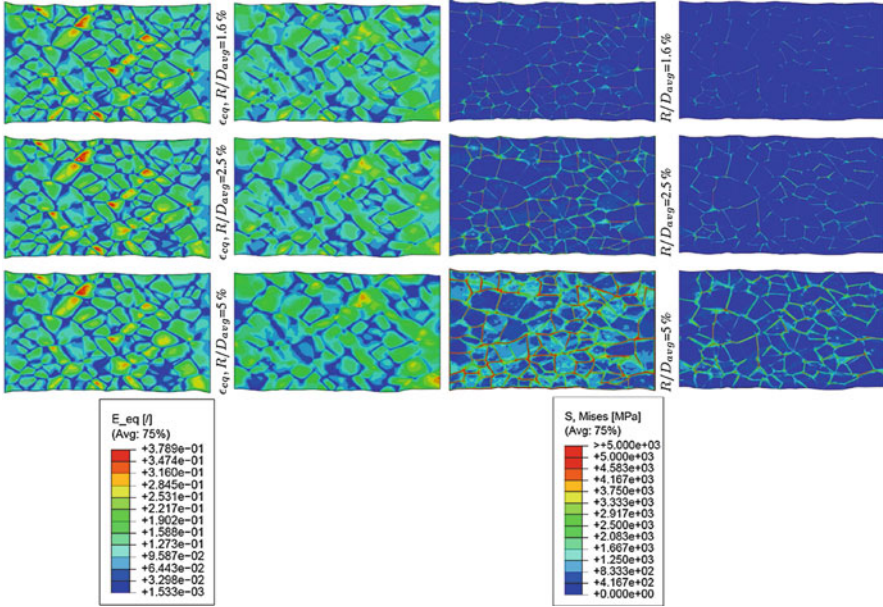


Fig. 4 The effect of hard (*left*) and soft (*right*) BC on the strain and stress fields in a 110 grain aggregate with $R/D_{avg} = 1.6\%$, 2.5% , and 5% and $\dot{\epsilon} = 0.2 \text{ s}^{-1}$

The same trend is observed for the two strain rates $\dot{\epsilon} = 0.2, 2 \text{ s}^{-1}$. Stiffer responses are obtained at higher strain rates which is in line with expectation. For soft BC higher R/D_{avg} ratios also lead to larger differences in the response between aggregates of different number of grains. So grain number influence is larger. For hard BC, the grain number influence on the macroscopic response is significantly smaller, while at the same time the macroscopic responses are stiffer compared to the soft BC responses. For a given R/D_{avg} ratio, both local responses of a polycrystalline aggregate are the same even though the grain sizes are different, resulting in the same macroscopic responses as well.

Hard and soft BCs have strong effect on the local stress/strain distribution as well. With hard BC the plastic slips at the grain boundaries are imposed to be zero, which constrains the deformations at the grain boundaries significantly. Consequently, the grain boundary stresses are higher; see Figs. 4 and 5 for two different strain rates where Mises equivalent stress and strains are plotted. Increasing the R/D_{avg} ratios increases the grain boundary stresses and widens the areas of higher stresses at the grain boundaries. This is especially evident for the hard BC but can also be clearly observed for the soft BC. In general, the grain boundary stress concentrations can be observed already at relative small R/D_{avg} ratio of 1.6% .

Both higher strain rates and R/D_{avg} ratios result in stiffer macroscopic responses, while R/D_{avg} significantly impacts the grain boundary stresses, where the influence of the rate on the grain boundary stress is considerably smaller.

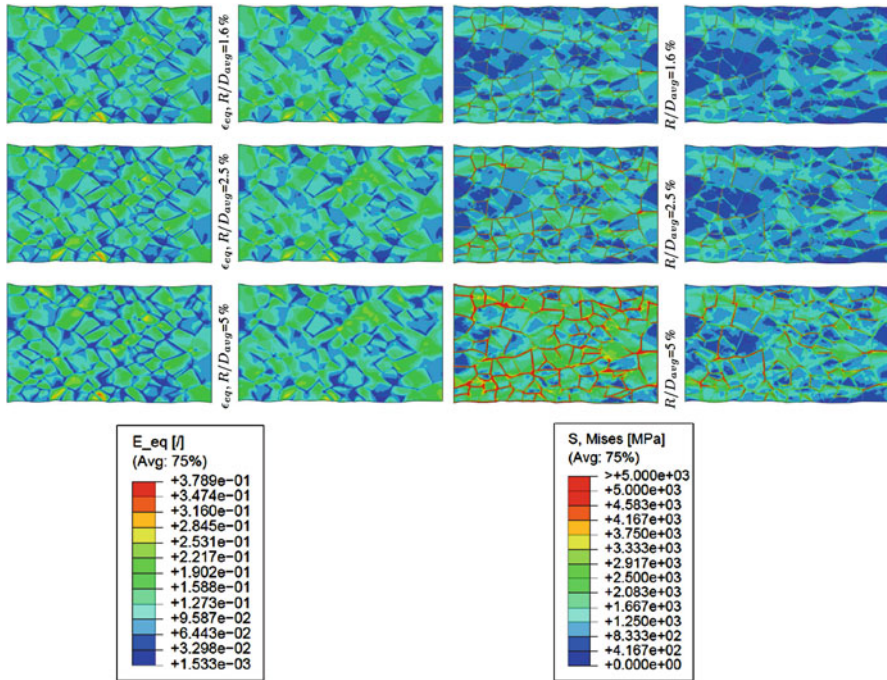


Fig. 5 The effect of hard (*left*) and soft (*right*) BC on the strain and stress fields in a 110 grain aggregate with $R/D_{avg} = 1.6\%$, 2.5% , and 5% and $\dot{\epsilon} = 2 \text{ s}^{-1}$

As for the material parameters, the rates used here are not related to a specific experiment, and they are quite high compared to 0.5 mm/mm/min (0.0083 s^{-1}) – 0.05 mm/mm/m (0.00083 s^{-1}) which ASTM E 8/E 8M-08 specifies for the strain rate during standard tensile testing of metallic materials ASTM 2009. Both material parameters and loading conditions should be identified for more realistic simulations.

Conclusion and Outlook

In this chapter, the thermodynamically consistent derivation of a higher-order, work-conjugate, plastic slip-based rate-dependent strain gradient crystal plasticity model is presented through both classical thermodynamics and rate variational formulation. After obtaining the coupled strong form of equations, the weak forms are derived through Galerkin method, and the complete finite element implementation is discussed in detail. The model presented here does not include any energetic hardening term. For the sake of simplicity and clarity, a linear dependence for the slip evolution equation is employed. Any extension of the model to include nonlinear terms could easily be done by following the presented steps above. The performance of the model is illustrated through polycrystalline examples using

Voronoi tessellation, and the influence of the boundary conditions, the loading rate, and the number and the size of the grains are studied. The current study considers only the hard and soft boundary conditions which are two extreme cases restricting the plastic slip completely or allowing the dislocation slip transfer without any resistance. Even though it is possible to capture the size effect due to the plasticity activity at the grain boundaries with the current nonlocal model, the boundary conditions are not completely physical, and the real behavior is somewhere in between these two situations. A grain boundary model is needed in order to incorporate the complex dislocation-grain boundary interaction mechanisms into the plasticity model. Therefore, the next chapter studies a particular grain boundary model and illustrates explicitly the effect of the misorientation between the grains and the orientation of the grain boundary. Due to its convex nature, the current model does not predict the intragranular plastic/dislocation microstructure formation and evolution. This requires the incorporation of an additional non-convex energy term.

Acknowledgements Tuncay Yalçinkaya gratefully acknowledges the support by the Scientific and Technological Research Council of Turkey (TÜBİTAK) under the 3001 Programme (Grant No. 215M381).

References

- A. Acharya, J.L. Bassani, Lattice incompatibility and a gradient theory of crystal plasticity. *J. Mech. Phys. Solids* **48**, 1565–1595 (2000)
- E.C. Aifantis, On the microstructural origin of certain inelastic models. *J. Eng. Mater. Technol.* **106**, 326–330 (1984)
- E.C. Aifantis, Strain gradient interpretation of size effects. *Int. J. Fract.* **95**, 299–314 (1999)
- K. Aifantis, J. Senger, D. Weygand, M. Zaiser, Discrete dislocation dynamics simulation and continuum modeling of plastic boundary layers in tricrystal micropillars. *IOP Conf. Ser. Mater. Sci. Eng.* **3**, 012025 (2009)
- A. Arsenlis, D.M. Parks, R. Becker, V.V. Bulatov, On the evolution of crystallographic dislocation density in non-homogeneously deforming crystals. *J. Mech. Phys. Solids* **52**, 1213–1246 (2004)
- M.F. Ashby, The deformation of plastically non-homogeneous materials. *Philos. Mag.* **21**, 399–424 (1970)
- ASTM, *Annual Book of ASTM Standards* (ASTM International, West Conshohocken, 2009)
- F. Aurenhammer, Voronoi diagrams – a survey of a fundamental geometric data structure. *ACM Comput. Surv.* **23**(3), 345–405 (1991)
- J.L. Bassani, Incompatibility and a simple gradient theory. *J. Mech. Phys. Solids* **49**, 1983–1996 (2001)
- E. Bayerschen, A.T. McBride, B.D. Reddy, T. Böhlke, Review on slip transmission criteria in experiments and crystal plasticity models. *J. Mater. Sci.* **51**(5), 2243–2258 (2016)
- C.J. Bayley, W.A.M. Brekelmans, M.G.D. Geers, A comparison of dislocation induced back stress formulations in strain gradient crystal plasticity. *Int. J. Solids Struct.* **43**, 7268–7286 (2006)
- P. van Beers, V. Kouznetsova, M. Geers, Defect redistribution within a continuum grain boundary plasticity model. *J. Mech. Phys. Solids* **83**, 243–262 (2015a)
- P. van Beers, V. Kouznetsova, M. Geers, Grain boundary interfacial plasticity with incorporation of internal structure and energy. *Mech. Mater.* **90**, 69–82 (2015b). Proceedings of the IUTAM Symposium on Micromechanics of Defects in Solids
- U. Borg, A strain gradient crystal plasticity analysis of grain size effects in polycrystals. *Eur. J. Mech. A-Solid.* **26**, 313–324 (2007)

- S.H. Chen, T.C. Wang, A new hardening law for strain gradient plasticity. *Acta Mater.* **48**, 3997–4005 (2000)
- A. Di Schino, J. Kenny, Grain size dependence of the fatigue behaviour of a ultrafine-grained AISI 304 stainless steel. *Mater. Lett.* **57**(21), 3182–3185 (2003)
- F.P.E. Dunne, D. Rugg, A. Walker, Lengthscale-dependent, elastically anisotropic, physically-based HCP crystal plasticity: application to cold-dwell fatigue in Ti alloys. *Int. J. Plast.* **23**, 1061–1083 (2007)
- L.P. Evers, W.A.M. Brekelmans, M.G.D. Geers, Non-local crystal plasticity model with intrinsic SSD and GND effects. *J. Mech. Phys. Solids* **52**, 2379–2401 (2004)
- X. Feaugas, H. Haddou, Grain-size effects on tensile behavior of nickel and AISI 316l stainless steel. *Metall. Mater. Trans. A* **34A**, 2329–2340 (2003)
- N.A. Fleck, J.W. Hutchinson, Strain gradient plasticity. *Adv. Appl. Mech.* **33**, 184–251 (1997)
- N.A. Fleck, J.W. Hutchinson, A reformulation of strain gradient plasticity. *J. Mech. Phys. Solids* **49**, 2245–2271 (2001)
- N.A. Fleck, G.M. Muller, M.F. Ashby, J.W. Hutchinson, Strain gradient plasticity: theory and experiment. *Acta Metall. Mater.* **42**, 475–487 (1994)
- N.A. Fleck, J.W. Hutchinson, J.R. Willis, Strain gradient plasticity under non-proportional loading. *Proc. R. Soc. A* **470**, 20140267 (2014)
- M.G.D. Geers, W.A.M. Brekelmans, C.J. Bayley, Second-order crystal plasticity: internal stress effects and cyclic loading. *Modell. Simul. Mater. Sci. Eng.* **15**, 133–145 (2007)
- D. Gottschalk, A. McBride, B. Reddy, A. Javili, P. Wriggers, C. Hirschberger, Computational and theoretical aspects of a grain-boundary model that accounts for grain misorientation and grain-boundary orientation. *Comput. Mater. Sci.* **111**, 443–459 (2016)
- P. Gudmundson, A unified treatment of strain gradient plasticity. *J. Mech. Phys. Solids* **52**, 1379–1406 (2004)
- M.E. Gurtin, On the plasticity of single crystals: free energy, microforces, plastic-strain gradients. *J. Mech. Phys. Solids* **48**, 989–1036 (2000)
- M.E. Gurtin, A gradient theory of single-crystal viscoplasticity that accounts for geometrically necessary dislocations. *J. Mech. Phys. Solids* **50**, 5–32 (2002)
- M.E. Gurtin, A finite-deformation, gradient theory of single-crystal plasticity with free energy dependent on densities of geometrically necessary dislocations. *Int. J. Plast.* **24**, 702–725 (2008)
- M.E. Gurtin, L. Anand, Thermodynamics applied to gradient theories involving the accumulated plastic strain: the theories of Aifantis and Fleck and Hutchinson and their generalization. *J. Mech. Phys. Solids* **57**, 405–421 (2009)
- C.S. Han, H. Gao, Y. Huang, W.D. Nix, Mechanism-based strain gradient crystal plasticity – I. Theory. *J. Mech. Phys. Solids* **53**, 1188–1203 (2005a)
- C.S. Han, H. Gao, Y. Huang, W.D. Nix, Mechanism-based strain gradient crystal plasticity – II. Analysis. *J. Mech. Phys. Solids* **53**, 1204–1222 (2005b)
- M.A. Haque, M.T.A. Saif, Strain gradient effect in nanoscale thin films. *Acta Mater.* **51**, 3053–3061 (2003)
- Y. Huang, S. Qu, K.C. Hwang, M. Li, H. Gao, A conventional theory of mechanism-based strain gradient plasticity. *Int. J. Plast.* **20**, 753–782 (2004)
- J.W. Hutchinson, Generalizing j_2 flow theory: fundamental issues in strain gradient plasticity. *Acta Mech. Sinica* **28**, 1078–1086 (2012)
- B. Klusemann, T. Yalçinkaya, Plastic deformation induced microstructure evolution through gradient enhanced crystal plasticity based on a non-convex helmholtz energy. *Int. J. Plast.* **48**, 168–188 (2013)
- B. Klusemann, T. Yalçinkaya, M.G.D. Geers, B. Svendsen, Application of non-convex rate dependent gradient plasticity to the modeling and simulation of inelastic microstructure development and inhomogeneous material behavior. *Comput. Mater. Sci.* **80**, 51–60 (2013)
- M. Kuroda, V. Tvergaard, On the formulations of higher-order strain gradient crystal plasticity models. *J. Mech. Phys. Solids* **56**, 1591–1608 (2008)
- G. Lancioni, T. Yalçinkaya, A. Cocks, Energy-based non-local plasticity models for deformation patterning, localization and fracture. *Proc. R. Soc. A* **471**: 20150275 (2015a)

- G. Lancioni, G. Zitti, T. Yalcinkaya, Rate-independent deformation patterning in crystal plasticity. *Key Eng. Mater.* **651–653**, 944–949 (2015b)
- V. Levkovitch, B. Svendsen, On the large-deformation- and continuum-based formulation of models for extended crystal plasticity. *Int. J. Solids Struct.* **43**, 7246–7267 (2006)
- L. Liang, F.P.E. Dunne, GND accumulation in bi-crystal deformation: crystal plasticity analysis and comparison with experiments. *Int. J. Mech. Sci.* **51**, 326–333 (2009)
- A. Ma, F. Roters, D. Raabe, A dislocation density based constitutive model for crystal plasticity FEM including geometrically necessary dislocations. *Acta Mater.* **54**, 2169–2179 (2006)
- H.B. Mühlhaus, E.C. Aifantis, A variational principle for gradient plasticity. *Int. J. Solids Struct.* **28**, 845–857 (1991)
- W. Nix, H. Gao, Indentation size effects in crystalline materials: a law for strain gradient plasticity. *J. Mech. Phys. Solids* **46**(3), 411–425 (1998)
- T. Ohashi, Crystal plasticity analysis of dislocation emission from micro voids. *Int. J. Plast.* **21**, 2071–2088 (2005)
- I. Özdemir, T. Yalçinkaya, Modeling of dislocation-grain boundary interactions in a strain gradient crystal plasticity framework. *Comput. Mech.* **54**, 255–268 (2014)
- A. Panteghini, L. Bardella, On the finite element implementation of higher-order gradient plasticity, with focus on theories based on plastic distortion incompatibility. *Comput. Methods Appl. Mech. Eng.* (2016). <https://doi.org/10.1016/j.cma.2016.07.045>
- P. Perzyna, Temperature and rate dependent theory of plasticity of crystalline solids. *Revue Phys. Appl.* **23**, 445–459 (1988)
- B.D. Reddy, The role of dissipation and defect energy in variational formulations of problems in strain-gradient plasticity. Part 1: polycrystalline plasticity. *Contin. Mech. Thermodyn.* **23**, 527–549 (2011a)
- B.D. Reddy, The role of dissipation and defect energy in variational formulations of problems in strain-gradient plasticity. Part 2: single-crystal plasticity. *Contin. Mech. Thermodyn.* **23**, 551–572 (2011b)
- J.R. Rice, Inelastic constitutive relations for solids: an internal variable theory and its application to metal plasticity. *J. Mech. Phys. Solids* **19**, 433–455 (1971)
- J.Y. Shu, N.A. Fleck, Strain gradient crystal plasticity: size-dependent deformation of bicrystals. *J. Mech. Phys. Solids* **47**, 297–324 (1999)
- M. Silhavy, *The Mechanics and Thermodynamics of Continuous Media*, 1st edn. (Springer, Berlin, 1997)
- J.S. Stölken, A.G. Evans, A microbend test method for measuring the plasticity length scale. *Acta Mater.* **46**, 5109–5115 (1998)
- B. Svendsen, On thermodynamic- and variational-based formulations of models for inelastic continua with internal length scales. *Comput. Methods Appl. Mech. Eng.* **193**, 5429–5452 (2004)
- B. Svendsen, S. Bargmann, On the continuum thermodynamic rate variational formulation of models for extended crystal plasticity at large deformation. *J. Mech. Phys. Solids* **58**, 1253–1271 (2010)
- J.G. Swadener, E.P. George, G.M. Pharra, The correlation of the indentation size effect measured with indenters of various shapes. *J. Mech. Phys. Solids* **50**, 681–694 (2002)
- G.I. Taylor, Plastic strain in metals. *J. Inst. Met.* **62**, 307–325 (1938)
- C.A. Volkert, E.T. Lilleodden, Size effects in the deformation of sub-micron columns. *Philos. Mag.* **86**, 5567–5579 (2006)
- G. Voyiadjis, R. Abu Al-Rub, Gradient plasticity theory with a variable length scale parameter. *Int. J. Solids Struct.* **42**(14), 3998–4029 (2005)
- J. Wang, J. Lian, J.R. Greer, W.D. Nix, K.S. Kim, Size effect in contact compression of nano- and microscale pyramid structures. *Acta Mater.* **54**, 3973–3982 (2006)
- T. Yalcinkaya, Microstructure evolution in crystal plasticity : strain path effects and dislocation slip patterning. PhD Thesis, Eindhoven University of Technology, 2011
- T. Yalçinkaya, Multi-scale modeling of microstructure evolution induced anisotropy in metals. *Key Eng. Mater.* **554–557**, 2388–2399 (2013)

-
- T. Yalcinkaya, G. Lancioni, Energy-based modeling of localization and necking in plasticity. *Proc. Mater. Sci.* **3**, 1618–1625 (2014)
- T. Yalcinkaya, W.A.M. Brekelmans, M.G.D. Geers, Deformation patterning driven by rate dependent nonconvex strain gradient plasticity. *J. Mech. Phys. Solids* **59**, 1–17 (2011)
- T. Yalcinkaya, W.A.M. Brekelmans, M.G.D. Geers, Non-convex rate dependent strain gradient crystal plasticity and deformation patterning. *Int. J. Solids Struct.* **49**, 2625–2636 (2012)
- S. Yefimov, I. Groma, E. van der Giessena, A comparison of a statistical-mechanics based plasticity model with discrete dislocation plasticity calculations. *J. Mech. Phys. Solids* **52**, 279–300 (2004)



Strain Gradient Crystal Plasticity: Intergranular Microstructure Formation

29

İzzet Özdemir and Tuncay Yalçinkaya

Contents

Introduction	1036
Strain Gradient Crystal Plasticity Framework and Grain Boundary Model	1039
Principle of Virtual Power: Macroscopic and Microscopic Energy Balances	1039
Free Energy Imbalance: Bulk Material	1042
Free Energy Imbalance: Interface	1045
Finite Element Implementation	1049
Numerical Examples	1052
Bi-crystal Specimen with Single Slip System	1053
Cylindrical Specimen with Three Slip System	1058
Conclusion and Outlook	1060
References	1062

Abstract

This chapter addresses the formation and evolution of inhomogeneous plastic deformation field between grains in polycrystalline metals by focusing on continuum scale modeling of dislocation-grain boundary interactions within a strain gradient crystal plasticity (SGCP) framework. Thermodynamically consistent extension of a particular strain gradient plasticity model, addressed previously

İ. Özdemir (✉)

Department of Civil Engineering, İzmir Institute of Technology, İzmir, Turkey

e-mail: izzetozdemir@iyte.edu.tr

T. Yalçinkaya

Aerospace Engineering Program, Middle East Technical University Northern Cyprus Campus, Guzelyurt, Mersin, Turkey

Department of Aerospace Engineering, Middle East Technical University, Ankara, Turkey

e-mail: yalcinka@metu.edu.tr

© Springer Nature Switzerland AG 2019

G. Z. Voyiadjis (ed.), *Handbook of Nonlocal Continuum Mechanics for Materials and Structures*, https://doi.org/10.1007/978-3-319-58729-5_4

1035

(see also, e.g., Yalcinkaya et al, *J Mech Phys Solids* 59:1–17, 2011), is presented which incorporates the effect of grain boundaries on plastic slip evolution explicitly. Among various choices, a potential-type non-dissipative grain boundary description in terms of grain boundary Burgers tensor (see, e.g., Gurtin, *J Mech Phys Solids* 56:640–662, 2008) is preferred since this is the essential descriptor to capture both the misorientation and grain boundary orientation effects. A mixed finite element formulation is used to discretize the problem in which both displacements and plastic slips are considered as primary variables. For the treatment of grain boundaries within the solution algorithm, an interface element is formulated. The capabilities of the framework is demonstrated through 3D bi-crystal and polycrystal examples, and potential extensions and currently pursued multi-scale modeling efforts are briefly discussed in the closure.

Keywords

Strain gradient plasticity · Grain boundary · Grain boundary-dislocation interaction · Misorientation · Grain boundary Burgers tensor

Introduction

At the grain scale, polycrystalline materials tend to develop heterogeneous plastic deformation fields due to variation of grain orientations, geometries, and defects. Grain boundaries are natural locations for plastic slip accumulation and geometrically necessary dislocations accommodating the gradients of inhomogeneous plastic strain. Grain boundaries have been subject to extensive research for decades due to their major influence on the mechanical properties of polycrystalline metals. Investigations on grain boundaries within polycrystalline materials date back to the observations of Hall and Petch. According to the so-called Hall-Petch effect, the yield strength of a polycrystalline metal specimen scales linearly with the inverse square root of the grain size. This is explained by the fact that grain boundaries behave as barriers against dislocation mobility, and they limit the mean free path of the dislocations thereby increasing strain hardening. Obviously, the grain size represents an intrinsic length scale of the polycrystalline system which inevitably influences the material response significantly. Conventional plasticity models do not incorporate a material length scale, and predictions of such theories involve only the lengths associated with the geometry of the whole solid body. In order to capture the grain size-dependent response, an internal length scale must be incorporated in the constitutive law. In the last two decades, the number of phenomenological and physically based strain gradient-type plasticity models has increased significantly. The intention of these models was to address the issue of internal length scale properly. Some of these models indeed consider the grain boundaries in a direct or smeared way; however understanding of dislocation-grain boundary interaction is far from being complete and constitutive models, and their numerical implementations are still subject of ongoing discussions. As stated by McDowell (2008), capturing the role of grain boundaries accurately within

continuum scale plasticity models is a challenging task. To this end, researchers have utilized different models and tools at different scales. In the context of grain boundary-dislocation interaction modeling, most of the existing efforts fall under one of the four categories identified as: (i) atomistic level studies, (ii) dislocation dynamics simulations, (iii) macroscale phenomenological modeling, and (iv) (multi-scale) crystal plasticity modeling.

When resolved at atomic scale, grain boundary is a thin transition layer from one grain (orientation) to the other, and it has its own structure. This structure endows the grain boundary with an initial energy (equilibrium energy) which is the focus of a considerable fraction of the literature on atomistic level grain boundary models; see, for example, Tschopp and McDowell (2007) and McDowell (2008). Although these studies mostly focus on specific grain boundaries (e.g., symmetric tilt boundaries) and a limited number of materials, the findings might be used effectively in multi-scale strain gradient plasticity type models. Developing a quantitative understanding through experiments on dislocation nucleation and slip transfer reactions at grain boundaries is extremely difficult. Therefore, atomistic models are the ideal tools to develop an understanding on unit mechanisms involved in dislocation absorption, transmission, and reflection within/from a grain boundary as exemplified by de Koning et al. (2002, 2003) and Spearot and McDowell (2009). However, significant challenges exist such as extending the number of slip transfer events to realistically large numbers, modeling nonequilibrium grain boundaries, and understanding dislocation activity in polycrystalline metals with heterogeneous compositions, e.g., impurities at grain boundaries or metallic alloys with heterophase interfaces.

Along with atomistic models, there have been a number of studies which have tried to embed grain boundaries within discrete dislocation dynamics framework (DDD, Van der Giessen and Needleman 1995) in a consistent way. At first, the focus was primarily on the treatment of slip transmission through grain boundaries within DDD. In this context, dislocation transmission rules based on transmission electron microscope (TEM) observations (see, e.g., Shen et al. 1986; Lee et al. 1989) were introduced into DDD codes (Li et al. 2009). As an alternative, the findings of atomistic level studies such as the one by de Koning et al. (2002) are adopted to introduce slip transmission criteria in DDD; see Kumar et al. (2010). However, the transmission of dislocations is not the only interaction mechanism to be considered. The absorption and emission of dislocations are also possible, but the rules for the incorporation of these phenomena are not clear yet.

The last two groups of studies mainly consist of strain gradient-type plasticity models taking into account the effect of inhomogeneous plastic strain distribution due to grain boundaries, i.e., high slip gradients in the vicinity of grain boundaries and related back-stress evolution, with either grain boundary conditions or with a decomposition of the grain into a core and a grain boundary effected zone. The majority of the strain gradient models could only treat the grain boundaries in two limiting cases as either impenetrable surfaces with no slip (hard boundaries) or slip without any resistance (soft boundaries). However, the proper interpretation of grain boundary effects could only be done by incorporating a grain boundary

description into strain gradient plasticity models which could cover other scenarios of dislocation-grain boundary interactions.

It was first recognized by Gudmundson (2004) that jump conditions across interfaces in strain gradient theories can be introduced and an additional contribution to the internal virtual work has to be considered along the interface between two plastically deforming phases. The interface conditions and traction expressions are derived, yet a polycrystal plasticity simulation was not performed. Later on Fredriksson and Gudmundson (2005) used this strain gradient framework with the interface model to capture dislocation built up at elastic/plastic interfaces. In a following work, Aifantis and Willis (2005) presented a similar approach which incorporates the physical properties of interfaces by means of interface potential appended to a phenomenological strain gradient plasticity model. To explore the capabilities of the model, Aifantis et al. (2006) calibrated the proposed interface potential through indentation tests, where the grain boundary “yield” reveals itself in the form of strain/displacement bursts observed in the load vs. tip displacement diagrams. In a parallel work, Abu Al-Rub (2008) extended a strain gradient plasticity framework by including an interface potential to investigate thin film mechanics. In the work of Massart and Pardoën (2010), interface elements were used to control/prescribe higher order boundary conditions stemming from the use of Fleck-Hutchinson gradient plasticity model (Fleck and Hutchinson 2001). The rate of plastic slip is initially set to zero at these interfaces and evolves with deformation mimicking the grain boundary relaxation mechanisms. Apart from these phenomenological gradient plasticity theories, there have been a number of attempts which have consolidated grain boundary effects within SGCP frameworks. One of the earliest ones was proposed by Ma et al. (2006) which incorporated the mechanical interaction between mobile dislocations and grain boundaries in a dislocation-based crystal plasticity model. The influence of grain boundaries is taken into account in terms of an additional energy affecting the dislocation velocity (plastic slip) in the bulk as opposed to the interfacial nature of grain boundaries. Nevertheless, the study reveals that the misorientation of the neighboring grains alone is not sufficient to describe the influence of grain boundaries. Borrowing the ideas of Aifantis and Willis (2005), Borg (2007) and Borg and Fleck (2007) incorporated an energy potential that penalizes crystallographic slip at grain boundaries into a SGCP theory for finite deformations. This approach does not take into account the misorientation of the neighboring grains or the grain boundary orientation, reflecting a rather phenomenological description. In a more recent study by Ekh et al. (2011), interface conditions sensitive to misorientation between grains were introduced in an ad hoc manner, where the slip resistance at grain boundaries decreases as the level of slip system alignment increases. However Gurtin (2008), which is also the point of departure for this chapter and has influenced many other recent works by van Beers et al. (2013), Özdemir and Yalçinkaya (2014), and Gottschalk et al. (2016), intrinsically includes the effect of both the misorientation between neighboring grains and the orientation of the grain boundary through grain boundary Burgers tensor.

Departing from the structure outlined by Gurtin (2008), in this chapter an extension of a rate-dependent strain gradient crystal plasticity model (convex

counterpart of Yalcinkaya et al. 2011 and Yalçinkaya et al. 2012) is presented in which grain boundary response is embedded in a thermodynamically consistent manner. Furthermore a mixed finite element-based solution algorithm is elaborated where both the plastic slips and displacement fields are taken as degrees of freedom.

The chapter is organized as follows. First, in section “[Strain Gradient Crystal Plasticity Framework and Grain Boundary Model](#),” the SGCP framework is recapitulated, and its consistent extension is presented. The particular form of the grain boundary potential is described in detail as well. In the following section, the finite element-based solution algorithm and the treatment of grain boundaries in a discrete setting are worked out. In section “[Numerical Examples](#),” the proposed formulation is assessed on the basis of three-dimensional examples. The chapter is closed by conclusion and outlook section, commenting on current efforts of extensions within a multi-scale modeling perspective.

Strain Gradient Crystal Plasticity Framework and Grain Boundary Model

In this section, within the context of continuum thermodynamics, crystal plasticity model is presented where the bulk behavior and the interface are treated separately. First, the force balances are derived via the principle of virtual power followed by the consideration of dissipation inequality for the bulk material and the interface successively.

Principle of Virtual Power: Macroscopic and Microscopic Energy Balances

In a geometrically linear kinematics setting, the time-dependent displacement field of a body with a grain boundary, as shown in Fig. 1, is denoted by $\mathbf{u} = \mathbf{u}(\mathbf{x}, t)$, where \mathbf{x} indicates the position of a material point. It has to be noted that the displacement field is continuous across the grain boundary. Furthermore, although the discussions are based on a body composed of two grains as depicted in Fig. 1, the following derivations are valid for bodies composed of multiple grains.

The strain tensor $\boldsymbol{\varepsilon}$ is defined as $\boldsymbol{\varepsilon} = \frac{1}{2}(\nabla\mathbf{u} + (\nabla\mathbf{u})^T)$, and the velocity vector is represented as $\mathbf{v} = \dot{\mathbf{u}}$. The strain is decomposed additively as

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}^e + \boldsymbol{\varepsilon}^p \quad (1)$$

into an elastic part $\boldsymbol{\varepsilon}^e$ and a plastic part $\boldsymbol{\varepsilon}^p$.

The plastic strain rate can be written as the summation of plastic slip rates on the individual slip systems, $\dot{\boldsymbol{\varepsilon}}^p = \sum_{\alpha} \dot{\gamma}^{\alpha} \mathbf{P}^{\alpha}$ with $\mathbf{P}^{\alpha} = \frac{1}{2}(\mathbf{s}^{\alpha} \otimes \mathbf{m}^{\alpha} + \mathbf{m}^{\alpha} \otimes \mathbf{s}^{\alpha})$ the symmetrized Schmid tensor, where \mathbf{s}^{α} and \mathbf{m}^{α} are the unit slip direction vector and unit normal vector on slip system α , respectively. The set

$$state = \boldsymbol{\varepsilon}^e, \gamma^{\alpha}, \nabla\gamma^{\alpha} \quad (2)$$

is chosen to be the state variables where γ^α and $\nabla\gamma^\alpha$ are the plastic slip and gradient of plastic slip on system α . Following the arguments of Gurtin (e.g., Gurtin 2000, 2002), the power expended by each independent rate-like kinematical descriptor is expressible in terms of an associated force consistent with its own balance. However, the basic kinematical fields of rate variables, namely, $\dot{\epsilon}^e$, \dot{u} , and $\dot{\gamma}^\alpha$ are not independent. It is therefore not immediately clear how the associated force balances are to be formulated, and, for that reason, these balances are established using the principle of virtual power.

The classical macroscopic system is defined by a traction $\mathbf{t}(\bar{\mathbf{n}})$ that expends power over the velocity $\dot{\mathbf{u}}$ and stress $\boldsymbol{\sigma}$ that expends power over the elastic strain rate $\dot{\boldsymbol{\epsilon}}^e$. There are no body forces acting on the system. The microscopic system is composed of a scalar microscopic stress π^α for each slip system that expends power over the slip rate $\dot{\gamma}^\alpha$, a vector of microscopic stress $\boldsymbol{\xi}^\alpha$ that expends power over the slip-rate gradient $\nabla\dot{\gamma}^\alpha$, a scalar microscopic traction $\chi^\alpha(\bar{\mathbf{n}})$ that expends power over $\dot{\gamma}^\alpha$, and a scalar interfacial microscopic grain boundary stress λ^α that expends power over the grain boundary slip rates $\dot{\gamma}_A^\alpha$ and $\dot{\gamma}_B^\alpha$ where approaching the grain boundary from different grains is designated by subscripts A and B. The virtual rates are collected in the generalized virtual velocity $\mathcal{V} = (\delta\dot{\mathbf{u}}, \delta\dot{\boldsymbol{\epsilon}}^e, \delta\dot{\gamma}^\alpha, \delta\dot{\gamma}_A^\alpha, \delta\dot{\gamma}_B^\alpha)$. The force systems are characterized through their work-conjugated nature with respect to the state variables. \mathcal{P}_{ext} is the power expended on the domain Ω , \mathcal{P}_{int} a concomitant expenditure of power within Ω stated as

$$\mathcal{P}_{ext}(\Omega, \mathcal{V}) = \int_S \mathbf{t}(\bar{\mathbf{n}}) \cdot \delta\dot{\mathbf{u}} \, dS + \sum_\alpha \int_S \chi^\alpha(\bar{\mathbf{n}}) \delta\dot{\gamma}^\alpha \, dS + \sum_\alpha \int_{S_{GB}} (\lambda_A^\alpha \delta\dot{\gamma}_A^\alpha + \lambda_B^\alpha \delta\dot{\gamma}_B^\alpha) \, dS \quad (3)$$

$$\mathcal{P}_{int}(\Omega, \mathcal{V}) = \int_\Omega \boldsymbol{\sigma} : \delta\dot{\boldsymbol{\epsilon}}^e \, d\Omega + \sum_\alpha \int_\Omega \pi^\alpha \delta\dot{\gamma}^\alpha \, d\Omega + \sum_\alpha \int_\Omega \boldsymbol{\xi}^\alpha \cdot \nabla\delta\dot{\gamma}^\alpha \, d\Omega$$

where λ_A and λ_B are the microscopic grain boundary stresses associated with the grain boundary of grain A and grain B, respectively. The boundary S is the union of outer boundaries of grains A and B; and S_{GB} is the grain boundary (the interface) as shown in Fig. 1.

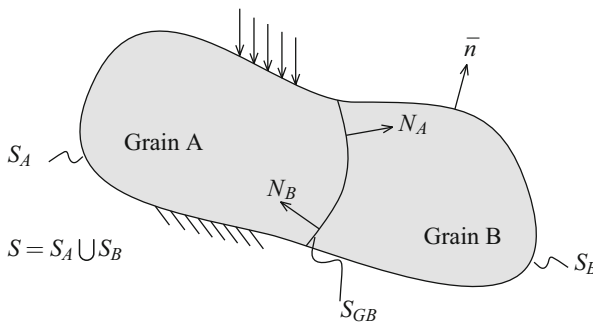


Fig. 1 A body with a grain boundary

Postulation of principle of virtual power states that given any set of virtual fields, the corresponding internal and external powers are balanced

$$\mathcal{P}_{ext}(\Omega, \mathcal{V}) = \mathcal{P}_{int}(\Omega, \mathcal{V}) \quad \forall \mathcal{V} \quad (4)$$

Now the consequences are derived following the restriction $\nabla \delta \dot{\mathbf{u}} = \delta \dot{\boldsymbol{\varepsilon}}^e + \delta \dot{\boldsymbol{\omega}}^e + \sum_{\alpha} \delta \dot{\gamma}^{\alpha} \mathbf{s}^{\alpha} \otimes \mathbf{m}^{\alpha}$. First a generalized virtual velocity without slip is considered, namely, $\delta \dot{\gamma}^{\alpha} = 0$ which means $\nabla \delta \dot{\mathbf{u}} = \delta \dot{\boldsymbol{\varepsilon}}^e + \delta \dot{\boldsymbol{\omega}}^e$. Considering that Cauchy stress is symmetric, the power balance becomes

$$\int_{\Omega} \boldsymbol{\sigma} : \nabla \delta \dot{\mathbf{u}} d\Omega = \int_S \mathbf{t}(\bar{\mathbf{n}}) \cdot \delta \dot{\mathbf{u}} dS \quad (5)$$

The conditions derived from equation (5) are well-known traction conditions

$$\mathbf{t}(\bar{\mathbf{n}}) = \boldsymbol{\sigma} \bar{\mathbf{n}} \quad (6)$$

and the classical linear momentum balance

$$\nabla \cdot \boldsymbol{\sigma} = \mathbf{0} \quad (7)$$

The microscopic counterparts of the conditions are obtained through a consideration of a generalized virtual velocity with $\delta \dot{\mathbf{u}} = \mathbf{0}$ with arbitrary $\delta \dot{\gamma}^{\alpha}$ field which results in

$$\sum_{\alpha} \delta \dot{\gamma}^{\alpha} (\mathbf{s}^{\alpha} \otimes \mathbf{m}^{\alpha}) = -\delta \dot{\boldsymbol{\varepsilon}}^e - \delta \dot{\boldsymbol{\omega}}^e \quad (8)$$

and the term $\boldsymbol{\sigma} : \delta \dot{\boldsymbol{\varepsilon}}^e$ becomes $\boldsymbol{\sigma} : \delta \dot{\boldsymbol{\varepsilon}}^e = -\boldsymbol{\sigma} : \sum_{\alpha} \delta \dot{\gamma}^{\alpha} (\mathbf{s}^{\alpha} \otimes \mathbf{m}^{\alpha}) - \boldsymbol{\sigma} : \delta \dot{\boldsymbol{\omega}}^e$. Using the symmetry of the Cauchy stress and the definition of the Schmid resolved stress τ^{α}

$$\boldsymbol{\sigma} : \delta \dot{\boldsymbol{\varepsilon}}^e = -\sum_{\alpha} \delta \dot{\gamma}^{\alpha} \tau^{\alpha} \quad (9)$$

For this case the power balance (4) is utilized again and the following form is obtained

$$\begin{aligned} & \sum_{\alpha} \int_S \chi^{\alpha}(\bar{\mathbf{n}}) \delta \dot{\gamma}^{\alpha} dS + \sum_{\alpha} \int_{S_{GB}} (\lambda_A^{\alpha} \delta \dot{\gamma}_A^{\alpha} + \lambda_B^{\alpha} \delta \dot{\gamma}_B^{\alpha}) dS = \\ & -\sum_{\alpha} \int_{\Omega} \tau^{\alpha} \delta \dot{\gamma}^{\alpha} d\Omega + \sum_{\alpha} \int_{\Omega} \pi^{\alpha} \delta \dot{\gamma}^{\alpha} d\Omega + \sum_{\alpha} \int_{\Omega} \boldsymbol{\xi}^{\alpha} \cdot \nabla \delta \dot{\gamma}^{\alpha} d\Omega \end{aligned} \quad (10)$$

which could be written as

$$\begin{aligned} & \sum_{\alpha} \int_{\Omega} (-\tau^{\alpha} + \pi^{\alpha} - \nabla \cdot \xi^{\alpha}) \delta \dot{\gamma}^{\alpha} d\Omega + \sum_{\alpha} \int_S (\xi^{\alpha} \cdot \bar{\mathbf{n}} - \chi^{\alpha}(\bar{\mathbf{n}})) \delta \dot{\gamma}^{\alpha} dS \\ & + \sum_{\alpha} \int_{S_{GB}} (\xi_A^{\alpha} \cdot \mathbf{N}_A - \lambda_A^{\alpha}) \delta \dot{\gamma}_A^{\alpha} dS + \sum_{\alpha} \int_{S_{GB}} (\xi_B^{\alpha} \cdot \mathbf{N}_B - \lambda_B^{\alpha}) \delta \dot{\gamma}_B^{\alpha} dS = 0 \end{aligned} \quad (11)$$

and has to be satisfied for all $\delta \dot{\gamma}^{\alpha}$. This argument yields the microscopic traction condition on the outer boundary of the bulk material

$$\chi^{\alpha}(\bar{\mathbf{n}}) = \xi^{\alpha} \cdot \bar{\mathbf{n}} \quad (12)$$

and the microscopic force balance inside the bulk material on each slip system α

$$\tau^{\alpha} - \pi^{\alpha} + \nabla \cdot \xi^{\alpha} = 0 \quad (13)$$

and microscopic grain boundary interface conditions on both side of the grain boundary

$$\begin{aligned} \xi_A^{\alpha} \cdot \mathbf{N}^A &= \lambda_A^{\alpha} \\ \xi_B^{\alpha} \cdot \mathbf{N}^B &= \lambda_B^{\alpha} \end{aligned} \quad (14)$$

Since $\mathbf{N}^A = -\mathbf{N}^B$, the interface conditions given by equation (14) can be expressed as

$$\begin{aligned} \lambda_A^{\alpha} &= \xi_A^{\alpha} \cdot \mathbf{N}^A \\ \lambda_B^{\alpha} &= -\xi_B^{\alpha} \cdot \mathbf{N}^A \end{aligned} \quad (15)$$

The details of the previous derivations for the case of single crystals (i.e., bulk material without GB terms) could be followed explicitly from Gurtin (2000, 2002) and Gurtin et al. (2007) and including the grain boundary terms presented in detail in Gurtin (2008).

Free Energy Imbalance: Bulk Material

The local internal power expression in the bulk material can be written as

$$P_i = \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}^e + \sum_{\alpha} (\pi^{\alpha} \dot{\gamma}^{\alpha} + \xi^{\alpha} \cdot \nabla \dot{\gamma}^{\alpha}) \quad (16)$$

and the local dissipation inequality results in

$$D = P_i - \dot{\psi} = \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}^e + \sum_{\alpha} (\pi^{\alpha} \dot{\gamma}^{\alpha} + \xi^{\alpha} \cdot \nabla \dot{\gamma}^{\alpha}) - \dot{\psi} \geq 0 \quad (17)$$

The material is assumed to be endowed with a free energy with different contributions according to

$$\psi(\mathbf{e}^e, \nabla\gamma^\alpha) = \psi_e(\mathbf{e}^e) + \psi_{\nabla\gamma}(\nabla\gamma^\alpha) \quad (18)$$

The free energy is assumed to depend only on the elastic strain and the dislocation density. In comparison to recent works of the authors (see, e.g., Yalcinkaya et al. 2011, Yalcinkaya et al. 2012, Yalcinkaya 2013, Yalcinkaya and Lancioni 2014, Lancioni et al. 2015a,b) where the framework is developed and used for modeling the inhomogeneous deformation field (microstructure) formation and evolution, the current chapter does not include the energetic hardening term ψ_γ in the bulk material, which could be convex or non-convex in nature. The details of such modeling approaches are discussed in one of the chapters here.

The time derivative of the free energy is expanded and equation (17) is elaborated to

$$\begin{aligned} D &= \boldsymbol{\sigma} : \dot{\mathbf{e}}^e + \sum_\alpha (\pi^\alpha \dot{\gamma}^\alpha + \boldsymbol{\xi}^\alpha \cdot \nabla \dot{\gamma}^\alpha - \frac{\partial \psi}{\partial \mathbf{e}^e} : \dot{\mathbf{e}}^e - \frac{\partial \psi}{\partial \gamma^\alpha} \dot{\gamma}^\alpha - \frac{\partial \psi}{\partial \nabla \gamma^\alpha} \cdot \nabla \dot{\gamma}^\alpha) \\ &= \underbrace{(\boldsymbol{\sigma} - \frac{d\psi_e}{d\mathbf{e}^e})}_0 : \dot{\mathbf{e}}^e + \sum_\alpha (\pi^\alpha) \dot{\gamma}^\alpha + \sum_\alpha \underbrace{(\boldsymbol{\xi}^\alpha - \frac{\partial \psi_{\nabla\gamma}}{\partial \nabla \gamma^\alpha})}_0 \cdot \nabla \dot{\gamma}^\alpha \geq 0 \end{aligned} \quad (19)$$

The stress $\boldsymbol{\sigma}$ and the microstress vectors $\boldsymbol{\xi}^\alpha$ are regarded as energetic quantities

$$\begin{aligned} \boldsymbol{\sigma} &= \frac{d\psi_e}{d\mathbf{e}^e} \\ \boldsymbol{\xi}^\alpha &= \frac{\partial \psi_{\nabla\gamma}}{\partial \nabla \gamma^\alpha} \end{aligned} \quad (20)$$

whereas π^α does have a dissipative contribution

$$D = \sum_\alpha (\pi^\alpha) \dot{\gamma}^\alpha \geq 0 \quad (21)$$

The multipliers of the plastic slip rates are identified as the set of dissipative stresses defined as $\sigma_{\text{dis}}^\alpha$

$$\sigma_{\text{dis}}^\alpha = \pi^\alpha \quad (22)$$

In order to satisfy the reduced dissipation inequality at the slip system level, the following constitutive equation is proposed

$$\sigma_{\text{dis}}^\alpha = \varphi^\alpha \text{sign}(\dot{\gamma}^\alpha) \quad (23)$$

where φ^α represents the mobilized slip resistance of the slip system under consideration

$$\varphi^\alpha = \frac{s^\alpha}{\dot{\gamma}_0^\alpha} |\dot{\gamma}^\alpha| \quad (24)$$

with s^α the resistance to dislocation slip which is assumed to be constant and with $\dot{\gamma}_0^\alpha$ a reference slip rate. Note that this relation assumes a linear relationship between strain rate and stress. In more quantitative simulations, small values of strain rate sensitivity exponent would be required. This choice was of course made for simplicity only, yet with a large similarity to discrete dislocation studies using linear drag relations. Simulations with other strain rate exponent do not change the qualitative nature of the examples, yet they will affect the rate-dependent (time-dependent) behavior. Substituting (24) into (23) gives

$$\dot{\gamma}^\alpha = \frac{\dot{\gamma}_0^\alpha}{s^\alpha} \sigma_{\text{dis}}^\alpha \quad (25)$$

which, upon substitution of $\sigma_{\text{dis}}^\alpha$ according to (22), takes the following form

$$\dot{\gamma}^\alpha = \frac{\dot{\gamma}_0^\alpha}{s^\alpha} (\pi^\alpha) \quad (26)$$

Finally, using the microforce balance (13) results in the plastic slip equation

$$\dot{\gamma}^\alpha = \frac{\dot{\gamma}_0^\alpha}{s^\alpha} (\tau^\alpha + \nabla \cdot \xi^\alpha) \quad (27)$$

Contributions of the free energies defined in (18) enter the slip equation via (20) with $\tau^\alpha = d\psi_e/d\boldsymbol{\varepsilon}^e : \mathbf{P}^\alpha$ and $\xi^\alpha = \partial\psi_{\nabla\gamma}/\partial\nabla\gamma^\alpha$. Quadratic forms are used for the elastic free energy ψ_e and the plastic slip gradients free energy contribution $\psi_{\nabla\gamma}$, i.e.,

$$\begin{aligned} \psi_e &= \frac{1}{2} \boldsymbol{\varepsilon}^e : {}^4\mathbf{C} : \boldsymbol{\varepsilon}^e \\ \psi_{\nabla\gamma} &= \sum_\alpha \frac{1}{2} A \nabla\gamma^\alpha \cdot \nabla\gamma^\alpha \end{aligned} \quad (28)$$

where ${}^4\mathbf{C}$ is the fourth-order elasticity tensor and A is a scalar quantity, which includes an internal length scale parameter, governing the effect of the plastic slip gradients on the internal stress field. It may be, e.g., expressed as $A = ER^2/(16(1-\nu^2))$ as used in Bayley et al. (2006) and Geers et al. (2007), where R physically represents the radius of the dislocation domain contributing to the internal stress field, ν is Poisson's ratio, and E is Young's modulus.

Free Energy Imbalance: Interface

Similarly, the dissipation associated with the grain boundary can be written as,

$$D_{GB} = P_{GB} - \dot{\psi}_{GB} \geq 0 \quad (29)$$

where

$$P_{GB} = \sum_{\alpha} \int_{S_{GB}} (\lambda_A^{\alpha} \dot{\gamma}_A^{\alpha} + \lambda_B^{\alpha} \dot{\gamma}_B^{\alpha}) dS \quad (30)$$

and ψ_{GB} is the free energy of the grain boundary which is essentially the coarse-grained representation of complex grain boundary-dislocation slip interaction mechanism such as dislocation transmission, emission, and dissociation of dislocations within the grain boundary. Transmission electron microscopy (TEM) studies and dislocation level models reveal that energetic and dissipative character of grain boundaries is dictated by the local geometrical features such as misorientation of grains and the grain boundary orientation. Therefore, an energy potential which is sensitive to noncoherent slip systems of neighboring grains and grain boundary orientation would reflect the underlying physics sufficiently well to a certain extent.

In order to account for essential geometrical features, the kinematic characterization presented in Gurtin (2008) is adopted in this work. In Gurtin (2008), the slip incompatibility of the neighboring grains is described in terms of the grain boundary Burgers tensor defined as

$$\mathbf{G} = \sum_{\alpha} [\gamma_B^{\alpha} \mathbf{s}_B^{\alpha} \otimes \mathbf{n}_B^{\alpha} - \gamma_A^{\alpha} \mathbf{s}_A^{\alpha} \otimes \mathbf{n}_A^{\alpha}] (\mathbf{N} \times) \quad (31)$$

where for any vector \mathbf{N} , $\mathbf{N} \times$ is the tensor with components $(\mathbf{N} \times)_{ij} = \varepsilon_{ikj} N_k$. In equation (31), the relative misorientation of grains is reflected by the difference term, and the grain boundary orientation is accounted for by the tensor $\mathbf{N} \times$. Furthermore, using (31), the magnitude of grain boundary Burgers tensor can be expressed in the following form

$$|\mathbf{G}|^2 = \sum_{\alpha} \sum_{\beta} \left(C_{AA}^{\alpha\beta} \gamma_A^{\alpha} \gamma_A^{\beta} + C_{BB}^{\alpha\beta} \gamma_B^{\alpha} \gamma_B^{\beta} + -2C_{AB}^{\alpha\beta} \gamma_A^{\alpha} \gamma_B^{\beta} \right) \quad (32)$$

in which the slip interaction moduli are introduced such that

$$\begin{aligned} C_{AA}^{\alpha\beta} &= \left(\mathbf{s}_A^{\alpha} \cdot \mathbf{s}_A^{\beta} \right) \left(\mathbf{n}_A^{\alpha} \times \mathbf{N} \right) \cdot \left(\mathbf{n}_A^{\beta} \times \mathbf{N} \right) \\ C_{AB}^{\alpha\beta} &= \left(\mathbf{s}_A^{\alpha} \cdot \mathbf{s}_B^{\beta} \right) \left(\mathbf{n}_A^{\alpha} \times \mathbf{N} \right) \cdot \left(\mathbf{n}_B^{\beta} \times \mathbf{N} \right) \\ C_{BB}^{\alpha\beta} &= \left(\mathbf{s}_B^{\alpha} \cdot \mathbf{s}_B^{\beta} \right) \left(\mathbf{n}_B^{\alpha} \times \mathbf{N} \right) \cdot \left(\mathbf{n}_B^{\beta} \times \mathbf{N} \right) \end{aligned} \quad (33)$$

Since $C_{AA}^{\alpha\beta}$ and $C_{BB}^{\alpha\beta}$ represent interactions between slip systems within grain A and grain B, respectively, they are called intra-grain interaction moduli, whereas $C_{AB}^{\alpha\beta}$ represent the interaction between slip systems of the two grains and called inter-grain interaction moduli.

Furthermore, by decomposing slip system normal vector $(\mathbf{n}_A^\alpha, \mathbf{n}_A^\beta, \mathbf{n}_B^\alpha, \mathbf{n}_B^\beta)$ into normal and tangential components with respect to grain boundary plane as

$$\mathbf{n}_A^\alpha = (\mathbf{n}_A^\alpha \cdot \mathbf{N}) \mathbf{N} + (\mathbf{I} - \mathbf{n}_A^\alpha \cdot \mathbf{N}) \mathbf{N} = \mathbf{n}_A^{\alpha,nor} + \mathbf{n}_A^{\alpha,tan} \quad (34)$$

and using $\varepsilon - \delta$ identity ($\varepsilon_{ijk}\varepsilon_{ipq} = \delta_{jp}\delta_{kq} - \delta_{jq}\delta_{kp}$), it can be shown that

$$(\mathbf{n}_A^\alpha \times \mathbf{N}) \cdot (\mathbf{n}_A^\beta \times \mathbf{N}) = \mathbf{n}_A^{\alpha,tan} \cdot \mathbf{n}_A^{\beta,tan} \quad (35)$$

Therefore, the interaction moduli can be expressed in the following alternative form

$$\begin{aligned} C_{AA}^{\alpha\beta} &= (\mathbf{s}_A^\alpha \cdot \mathbf{s}_A^\beta) (\mathbf{n}_A^{\alpha,tan} \cdot \mathbf{n}_A^{\beta,tan}) \\ C_{AB}^{\alpha\beta} &= (\mathbf{s}_A^\alpha \cdot \mathbf{s}_B^\beta) (\mathbf{n}_A^{\alpha,tan} \cdot \mathbf{n}_B^{\beta,tan}) \\ C_{BB}^{\alpha\beta} &= (\mathbf{s}_B^\alpha \cdot \mathbf{s}_B^\beta) (\mathbf{n}_B^{\alpha,tan} \cdot \mathbf{n}_B^{\beta,tan}) \end{aligned} \quad (36)$$

which allows a clear geometric interpretation as illustrated in Fig. 2. For the sake of clarity, a two-dimensional bi-crystal with slip system α located in grain A and slip system β located in grain B is considered, and different geometric configurations are depicted in Fig. 2.

For the intra-grain moduli ($C_{AA}^{\alpha\alpha}$) and ($C_{BB}^{\beta\beta}$), the first inner product yields unity, and the second inner product term measures the contribution of this slip system to the slip incompatibility depending on the orientation of slip system with respect to the grain boundary orientation. If the slip system orientation coincides with the orientation of the grain boundary, in other words if the slip system is parallel to the grain boundary, then there is no interaction between the slip system and the grain boundary; therefore the slip system does not contribute to slip incompatibility.

In the inter-grain interaction moduli ($C_{AA}^{\alpha\alpha}$) expression, the first inner product term is the measure of geometric coherency of the slip systems, taking the maximum value of unity when the systems (slip system α in grain A and slip system β in grain B) are parallel. Similar to the intra-grain moduli, the second inner product term is the influence of the orientation of slip systems α and β with respect to the grain boundary orientation.

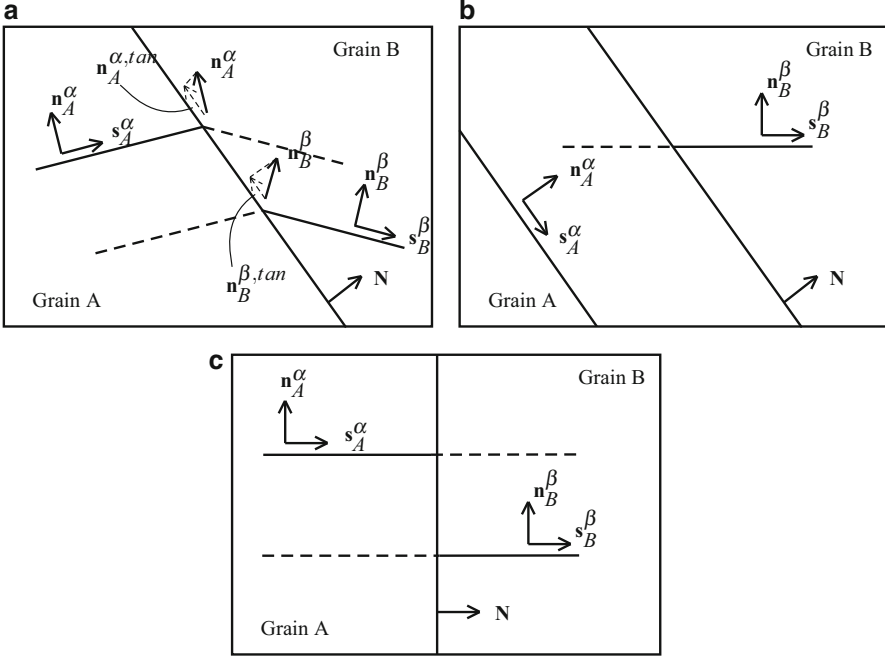


Fig. 2 (a) An arbitrary configuration, (b) an inclined grain boundary and a parallel slip system, and (c) coherent slip systems with a vertical grain boundary

At this stage ignoring the dissipative effects, a simple free energy potential of the form

$$\psi_{GB} = \frac{1}{2} \kappa |\mathbf{G}|^2 \tag{37}$$

is proposed where κ is a positive constant modulus. With this particular form in hand (32), rate of free energy can be expressed as

$$\dot{\psi}_{GB} = \kappa \sum_{\alpha} \sum_{\beta} \left[\left(\gamma_B^{\beta} C_{BB}^{\alpha\beta} - \gamma_A^{\beta} C_{BA}^{\alpha\beta} \right) \dot{\gamma}_B^{\alpha} - \left(\gamma_B^{\beta} C_{AB}^{\alpha\beta} - \gamma_A^{\beta} C_{AA}^{\alpha\beta} \right) \dot{\gamma}_A^{\alpha} \right] \tag{38}$$

Then the dissipation inequality reads as

$$\left[\lambda_A^{\alpha} - \kappa \sum_{\beta} \left(\gamma_A^{\beta} C_{AA}^{\alpha\beta} - \gamma_B^{\beta} C_{AB}^{\alpha\beta} \right) \right] \dot{\gamma}_A^{\alpha} + \left[\lambda_B^{\alpha} - \kappa \sum_{\beta} \left(\gamma_B^{\beta} C_{BB}^{\alpha\beta} - \gamma_A^{\beta} C_{BA}^{\alpha\beta} \right) \right] \dot{\gamma}_B^{\alpha} \geq 0 \tag{39}$$

Since a non-dissipative grain boundary response is assumed, equation (39) leads to following identifications

$$\begin{aligned}\lambda_A^\alpha &= \kappa \sum_\beta \left(\gamma_A^\beta C_{AA}^{\alpha\beta} - \gamma_B^\beta C_{AB}^{\alpha\beta} \right) \\ \lambda_B^\alpha &= \kappa \sum_\beta \left(\gamma_B^\beta C_{BB}^{\alpha\beta} - \gamma_A^\beta C_{BA}^{\alpha\beta} \right)\end{aligned}\quad (40)$$

In case of single slip system (α) in grain A and single slip system (β) in grain B, for an arbitrary orientation other than the orientation where one of the slip systems (α or β) is parallel to the grain boundary, the interaction moduli read as

$$\begin{aligned}C_{AA}^{\alpha\alpha} &= \mathbf{n}_A^{\alpha,tan} \cdot \mathbf{n}_A^{\alpha,tan} \\ C_{BB}^{\beta\beta} &= \mathbf{n}_B^{\beta,tan} \cdot \mathbf{n}_B^{\beta,tan} \\ C_{AB}^{\alpha\beta} &= \left(\mathbf{s}_A^\alpha \cdot \mathbf{s}_B^\beta \right) \left(\mathbf{n}_A^{\alpha,tan} \cdot \mathbf{n}_A^{\beta,tan} \right)\end{aligned}\quad (41)$$

In case of coherent slip systems, the inter-grain interaction moduli reduce to

$$C_{AB}^{\alpha\beta} = \mathbf{n}_A^{\alpha,tan} \cdot \mathbf{n}_B^{\beta,tan} \quad (42)$$

which implies

$$\begin{aligned}\lambda_A^\alpha &= \bar{\kappa} \left(\gamma_A^\alpha - \gamma_B^\beta \right) \\ \lambda_B^\beta &= \bar{\kappa} \left(\gamma_B^\beta - \gamma_A^\alpha \right)\end{aligned}\quad (43)$$

where $\bar{\kappa} = \left(\mathbf{n}_A^{\alpha,tan} \cdot \mathbf{n}_B^{\beta,tan} \right) \kappa$. If the slip systems are coherent and the grain boundary is perpendicular to these parallel slip systems, then the moduli take the following values

$$C_{AA}^{\alpha\alpha} = 1.0, C_{BB}^{\beta\beta} = 1.0, C_{AB}^{\alpha\alpha} = 1.0 \quad (44)$$

implying that

$$\begin{aligned}\lambda_A^\alpha &= \kappa \left(\gamma_A^\alpha - \gamma_B^\alpha \right) \\ \lambda_B^\alpha &= \kappa \left(\gamma_B^\alpha - \gamma_A^\alpha \right)\end{aligned}\quad (45)$$

The difference between the slips is being penalized by the grain boundary strength κ . When the slip system α of grain A is parallel to the grain boundary, the interaction moduli simplify to

$$C_{AA}^{\alpha\alpha} = 0, C_{AB}^{\alpha\beta} = 0, C_{BB}^{\beta\beta} = 1.0 \quad (46)$$

yielding

$$\lambda_A = 0, \lambda_B = \kappa \gamma_B^\beta \quad (47)$$

Since, in this configuration, grain A is not interacting with the grain boundary and grain B, the corresponding interaction moduli turn out to be zero. Depending on the value of κ (“the strength of the interface”), the grain boundary may act as a soft ($\kappa = 0$) or hard ($\kappa = \infty$) boundary for grain B as reflected by equation (47).

Finite Element Implementation

In order to solve the initial boundary value problem for the grain boundary enhanced strain gradient crystal plasticity framework, a mixed finite element formulation is used. The displacement \mathbf{u} and plastic slips γ^α are taken as primary variables, and these fields are determined within the problem domain by solving simultaneously the linear momentum balance and the microscopic force balance. Therefore the strong form of these equations reads as

$$\begin{aligned} \nabla \cdot \boldsymbol{\sigma} &= \mathbf{0} \\ \dot{\gamma}^\alpha - \frac{\dot{\gamma}_0^\alpha}{s^\alpha} \tau^\alpha - \frac{\dot{\gamma}_0^\alpha}{s^\alpha} \nabla \cdot \boldsymbol{\xi}^\alpha &= 0 \end{aligned} \quad (48)$$

complemented with the boundary conditions (6), (12) on the associated outer boundaries and the interface conditions (14) on the grain boundary, respectively. It has to be noted that the slip evolution equation is in fact the microforce balance (equation (13)) reexpressed in an alternative form with the aid of equations (22), (23), and (24). The weak forms of the balance equations are obtained through multiplication by weighting functions $\delta \mathbf{u}$ and $\delta \gamma^\alpha$ and integration over the domain Ω , yielding

$$\begin{aligned} G_u &= \int_\Omega \nabla \delta \mathbf{u} : \boldsymbol{\sigma} d\Omega - \int_S \delta \mathbf{u} \cdot \mathbf{t} dS \\ G_\gamma^\alpha &= \int_\Omega \delta \gamma^\alpha \dot{\gamma}^\alpha d\Omega - \int_\Omega \frac{\dot{\gamma}_0^\alpha}{s^\alpha} \delta \gamma^\alpha \tau^\alpha d\Omega + \int_\Omega \frac{\dot{\gamma}_0^\alpha}{s^\alpha} \nabla \delta \gamma^\alpha \cdot A \nabla \gamma^\alpha d\Omega \\ &\quad - \int_S \frac{\dot{\gamma}_0^\alpha}{s^\alpha} \delta \gamma^\alpha \chi^\alpha dS - \int_{S_{GB}} \frac{\dot{\gamma}_0^\alpha}{s^\alpha} \delta \gamma_A^\alpha \lambda_A^\alpha dS - \int_{S_{GB}} \frac{\dot{\gamma}_0^\alpha}{s^\alpha} \delta \gamma_B^\alpha \lambda_B^\alpha dS \end{aligned} \quad (49)$$

where the grain boundary contributions appear as additional terms in the weak form of microstress balance. Furthermore, \mathbf{t} in equation (49) is the external traction vector on the boundary S , and $\chi^\alpha = A \nabla \gamma^\alpha \cdot \bar{\mathbf{n}}$. Using a standard Galerkin approach,

the unknown fields of displacements, slips, and associated weighting functions are interpolated via

$$\begin{aligned} \delta \mathbf{u} &= \underline{N}^u \delta \tilde{\mathbf{u}} & \mathbf{u} &= \underline{N}^u \tilde{\mathbf{u}} \\ \delta \gamma^\alpha &= \underline{N}^\gamma \delta \tilde{\gamma}^\alpha & \gamma^\alpha &= \underline{N}^\gamma \tilde{\gamma}^\alpha \end{aligned} \tag{50}$$

where $\tilde{\mathbf{u}}$, $\delta \tilde{\mathbf{u}}$, $\tilde{\gamma}^\alpha$, and $\delta \tilde{\gamma}^\alpha$ are the columns containing the nodal variables of a particular element. Adopting a discretization by ten-noded tetrahedra elements, quadratic interpolation for the displacement field and linear interpolation for the slips are used. Referring to Fig. 3, for a ten-noded tetrahedra, only the corner nodes have the slip degrees of freedom, whereas all nodes have displacement degrees of freedom.

To facilitate the integration of the grain boundary contributions, 12-noded zero thickness interface elements are used which are inserted along the grain boundaries; please see Fig. 3. By means of these elements, one has the access to the slip values along the grain boundary as approached from grain A and grain B. However, it is important to note that the interface elements do not possess any kind of mechanical cohesive behavior and do not cause discontinuity in displacement field. In the solution phase, the displacement continuity across the grain boundary is fulfilled by means of equality constraints (rigid ties) enforcing the same dis-

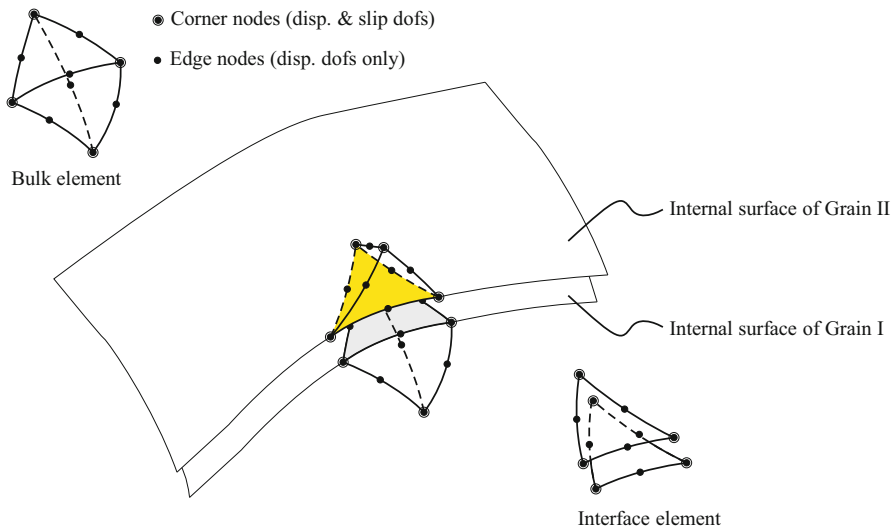


Fig. 3 Element types used for discretization. For ease of illustration, grain boundary and interface element are shown with a gap

placement field for the corresponding nodes of on the two sides of an interface element.

With a 12-noded interface element, $\delta\gamma_A^\alpha$ and $\delta\gamma_B^\alpha$ can be expressed as

$$\begin{bmatrix} \delta\gamma_A^\alpha \\ \delta\gamma_B^\alpha \end{bmatrix} = \begin{bmatrix} N_1 & N_2 & N_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & N_1 & N_2 & N_3 \end{bmatrix} \begin{bmatrix} \delta\gamma_1^\alpha \\ \delta\gamma_2^\alpha \\ \delta\gamma_3^\alpha \\ \delta\gamma_4^\alpha \\ \delta\gamma_5^\alpha \\ \delta\gamma_6^\alpha \end{bmatrix} = \underline{N}_{GB} \underline{\delta\gamma}_{GB}^\alpha \quad (51)$$

in terms of nodal virtual slip variables $\delta\gamma_i$ and the standard shape functions (N_1, N_2, N_3) for a linear triangle. In the columns of nodal quantities such as $\underline{\delta\gamma}_{GB}^\alpha$, subscripts designate the node number, and the superscripts are reserved for slip system identity. Introducing $\underline{\lambda} = [\lambda_A^\alpha \quad \lambda_B^\alpha]^T$, the grain boundary integral for the particular slip system α can be expressed as

$$\int_{S_{GB}} \frac{\dot{\gamma}_0^\alpha}{s^\alpha} (\lambda_A^\alpha \delta\gamma_A^\alpha + \lambda_B^\alpha \delta\gamma_B^\alpha) dS = \underline{\delta\gamma}_\alpha^T \int_{S_{GB}} \frac{\dot{\gamma}_0^\alpha}{s^\alpha} \underline{N}_{GB}^T \underline{\lambda} dS \quad (52)$$

Introducing the $2n \times 6n$ \underline{T} matrix (where n is the number of slip systems)

$$\underline{T} = \begin{bmatrix} N_1 & N_2 & N_3 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & N_1 & N_2 & N_3 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & N_1 & N_2 & N_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & N_1 & N_2 & N_3 \end{bmatrix} \quad (53)$$

and the rows

$$\begin{aligned} \underline{C}_A &= [C_{AA}^{\alpha 1} \quad -C_{AB}^{\alpha 1} \quad C_{AA}^{\alpha 2} \quad -C_{AB}^{\alpha 2} \quad \dots \quad C_{AA}^{\alpha n} \quad -C_{AB}^{\alpha n}] \\ \underline{C}_B &= [-C_{BA}^{\alpha 1} \quad C_{BB}^{\alpha 1} \quad -C_{BA}^{\alpha 2} \quad C_{BB}^{\alpha 2} \quad \dots \quad -C_{BA}^{\alpha n} \quad C_{BB}^{\alpha n}] \\ \underline{\gamma}^T &= [\gamma_1^n \quad \gamma_2^n \quad \gamma_3^n \quad \gamma_4^n \quad \gamma_5^n \quad \gamma_6^n \quad \dots \quad \gamma_1^n \quad \gamma_2^n \quad \gamma_3^n \quad \gamma_4^n \quad \gamma_5^n \quad \gamma_6^n] \end{aligned} \quad (54)$$

for slip system α , the interface element contribution reads as

$$G_{GB,\gamma}^{\alpha,e} = -\underline{\delta\gamma}_\alpha^T \int_{S_{GB}^e} \underline{N}_{GB}^T \underline{D} \underline{\gamma} dS^e \quad (55)$$

where

$$\underline{D} = \kappa \frac{\dot{\gamma}_0^\alpha}{s^\alpha} \begin{bmatrix} \underline{C}_A \underline{T} \\ \underline{C}_B \underline{T} \end{bmatrix} \tag{56}$$

By employing backward Euler time integration for $\dot{\gamma}^\alpha$, one can write $\dot{\gamma}^\alpha = [\gamma_{n+1}^\alpha - \gamma_n^\alpha]/\Delta t$ with $\Delta t = t_{n+1} - t_n$, and fully discrete weak forms for element e read

$$\begin{aligned} G_u^e &= \delta \underline{u}^T \left[\int_{\Omega^e} \underline{B}^u \underline{\sigma} d\Omega^e - \int_{S^e} \underline{N}^u \underline{t} dS^e \right] \\ G_{\gamma^{\alpha,e}} &= \delta \tilde{\gamma}^\alpha \left[\int_{\Omega^e} \underline{N}^{\gamma T} \underline{N}^\gamma \left[\frac{\gamma_{n+1}^\alpha - \gamma_n^\alpha}{\Delta t} \right] d\Omega^e - \int_{\Omega^e} \frac{\dot{\gamma}_0^\alpha}{s^\alpha} \underline{N}^{\gamma T} \underline{\tau}^\alpha d\Omega^e \right] \\ &\quad + \delta \tilde{\gamma}^\alpha \left[\int_{\Omega^e} \frac{\dot{\gamma}_0^\alpha}{s^\alpha} A \underline{B}^{\gamma T} \underline{B}^\gamma \underline{\gamma}^\alpha d\Omega^e - \int_{S^e} \frac{\dot{\gamma}_0^\alpha}{s^\alpha} \underline{N}^{\gamma T} \underline{\chi}^\alpha dS^e \right] \end{aligned} \tag{57}$$

By adding the integrals for the interface and bulk discretizations ((56) and (57), respectively) and enforcing $\sum_{i=1}^{n_{el}} G_{GB,\gamma}^{\alpha,e} + \sum_{i=1}^{n_{el}} G_\gamma^{\alpha,e} = 0$, the weak form for the whole domain is obtained. The weak forms of the balance equations are linearized with respect to the increments of the primary variables \underline{u} and γ^α and solved by means of a Newton-Raphson solution procedure for the corrective terms $\Delta \underline{u}$ and $\Delta \gamma^\alpha$. The resulting system of linear equations can be written in the following compact form

$$\begin{bmatrix} \underline{K}^{uu} & \underline{K}^{u\gamma} \\ \underline{K}^{\gamma u} & \underline{K}^{\gamma\gamma} + \underline{K}_{GB}^{\gamma\gamma} \end{bmatrix} \begin{bmatrix} \Delta \underline{u} \\ \Delta \gamma^\alpha \end{bmatrix} = \begin{bmatrix} -\underline{R}^u + \underline{R}_u^{ext} \\ -\underline{R}^\gamma + \underline{R}_\gamma^{ext} \end{bmatrix} \tag{58}$$

where \underline{K}^{uu} , $\underline{K}^{u\gamma}$, $\underline{K}^{\gamma u}$, and $\underline{K}^{\gamma\gamma}$ represent the global tangent matrices of the bulk. $\underline{K}_{GB}^{\gamma\gamma}$ is the contribution of the interface elements, while \underline{R}^u and \underline{R}^γ are the global residual columns. The contributions \underline{R}_u^{ext} and $\underline{R}_\gamma^{ext}$ originate from the boundary terms.

Numerical Examples

The presented framework is implemented both for 2D and 3D problems in Abaqus V 6.12-1 through user element capabilities. Since the 2D implementation and corresponding examples were documented in Özdemir and Yalçinkaya (2014), in this work the focus is on 3D formulation. To this end, a ten-noded tetrahedron with quadratic displacement and linear slip interpolations is developed and used for the discretization of the grain interiors. The slip degrees of freedom are defined on the corner nodes, whereas the edge nodes have only displacement degrees of freedom. For the treatment of grain boundaries, 12-noded zero thickness interface

element is implemented by following the formulation presented in section “[Finite Element Implementation](#).” Although the interface element has 12 nodes, only the corner nodes and the associated slip degrees of freedom are used for interpolation (see Fig. 3). As mentioned before, the displacement continuity is achieved by using equality constraints (rigid links) enforcing the same displacements on the two sides of a grain boundary.

Bi-crystal Specimen with Single Slip System

To assess the basic characteristics of the proposed formulation, a relatively simple geometry of a bi-crystal specimen is considered. Mechanical and slip boundary conditions are schematically shown in Fig. 4, and the material parameters are given in Table 1. The front face is displaced with a constant rate in positive z -direction by $10\ \mu\text{m}$ in 0.5 s. For the sake of clarity, a single slip system with $\vec{m} = (0, \sqrt{2}/2, \sqrt{2}/2)$ and $\vec{s} = (0, -\sqrt{2}/2, \sqrt{2}/2)$ is considered. In fact, this is a slip system solely defined in YZ plane and making an angle of 45° with the positive Z -axis. As far as grain boundary (GB) orientation is concerned, two different cases, namely, a vertical GB and a tilted one, are treated separately; please see Fig. 4. In what follows, slip profiles along the specimen are plotted with reference to the longitudinal path (from the front face to the back face of the specimen) shown in Fig. 4.

To start with, the case where the GB is vertical and slip systems of the two grains are perfectly coherent is considered. In other words, the misorientation between the

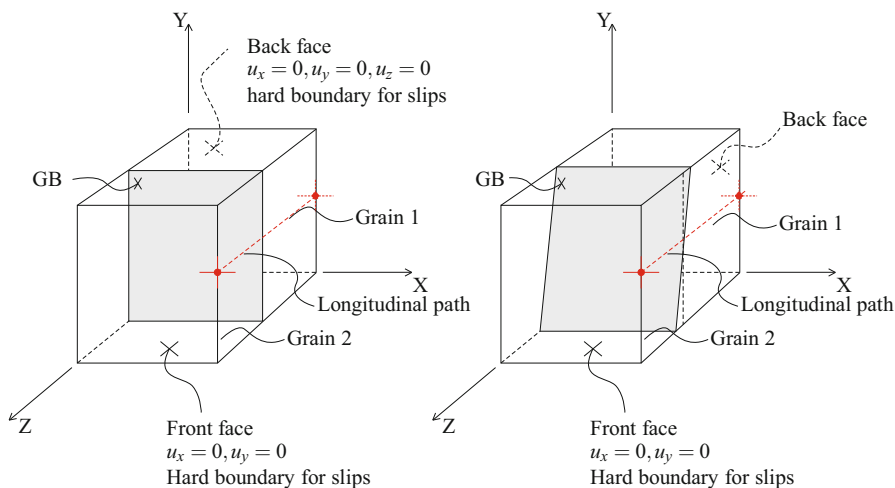


Fig. 4 Bi-crystal example with a vertical (*left*) and a tilted (*right*) grain boundary. Each grain is cubic and 50 by 50 by 50 μm in dimensions

Table 1 Geometry and material parameters for example 4.1 and example 4.2

	Example 4.1	Example 4.2
E (MPa)	70,000.0	70,000.0
ν (-)	0.33	0.33
$\dot{\gamma}_0$ (s^{-1})	0.0115	0.0115
s (MPa)	20	20
r (mm)	0.005	0.005
\vec{m}_1	$(0, \sqrt{2}/2, \sqrt{2}/2)$	$(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$
\vec{s}_1	$(0, -\sqrt{2}/2, \sqrt{2}/2)$	$(-\sqrt{2}/2, \sqrt{2}/2, 0)$
\vec{m}_2	-	$(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$
\vec{s}_2	-	$(0, \sqrt{2}/2, -\sqrt{2}/2)$
\vec{m}_3	-	$(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$
\vec{s}_3	-	$(\sqrt{2}/2, 0, -\sqrt{2}/2)$

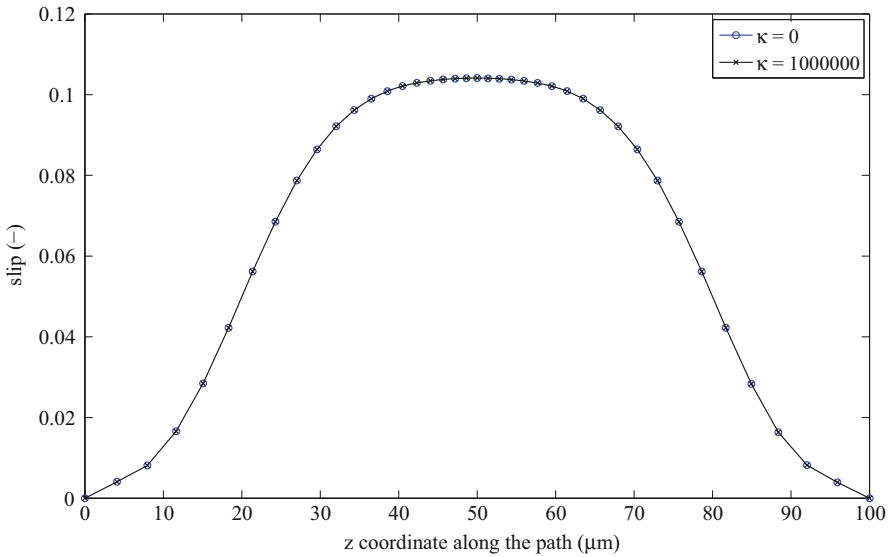


Fig. 5 Slip distribution along the longitudinal path for two different κ values

two grains is zero. For this case, two analyses with different GB strength (κ) are conducted. The slip distributions along the longitudinal path are shown in Fig. 5.

$\kappa = 0$ corresponds to a bi-crystal without a grain boundary, and $\kappa = 1,000,000$ is very close to a hard boundary as is going to be shown shortly. In fact the two solutions are identical which are consistent with the finding of section “Finite Element Implementation” stating that when the misorientation is zero, the grains do not “feel” the existence of a grain boundary. According to the analysis conducted in section “Finite Element Implementation,” the grains should not “feel” even an inclined grain boundary as long as the slip systems of the two grains are aligned. In order to test this, a set of analysis with crystallographically aligned bi-crystal with a

tilted GB (by 30°, see Fig. 4) is conducted. In Fig. 6, the slip contours within the bi-crystal specimen cut by a vertical GB and a tilted GB are depicted. For both cases, $\kappa = 1,000,000$ and from Fig. 5, it is known that for a vertical GB with aligned grains, the response is insensitive to κ . Identical contour seen in Fig. 6 verifies that an inclined grain boundary does not have any influence on the slip distribution provided that there is no misorientation.

In the next stage, the problem is slightly modified by rotating grain 2 around x axis by 15°. In this configuration, there is a misalignment between the slip systems of the two grains, and the slip distribution along the longitudinal path for different κ values are shown in Fig. 7. The influence of the grain boundary is clearly visible, and as the interface strength is increased, the slip distribution in the vicinity of the

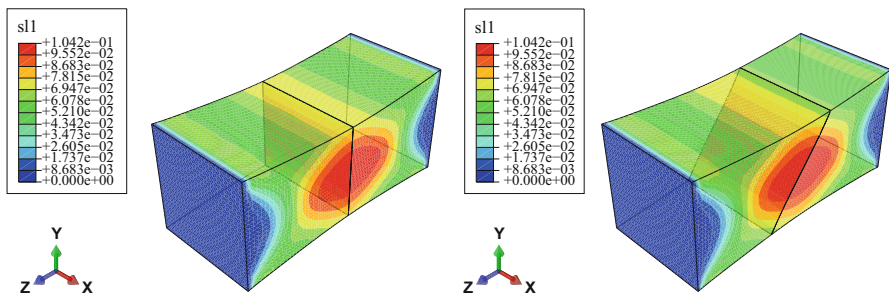


Fig. 6 Slip contours for a flat GB (top) and for a tilted GB. For both cases the misorientation angle between the grains is zero and $\kappa = 1,000,000$

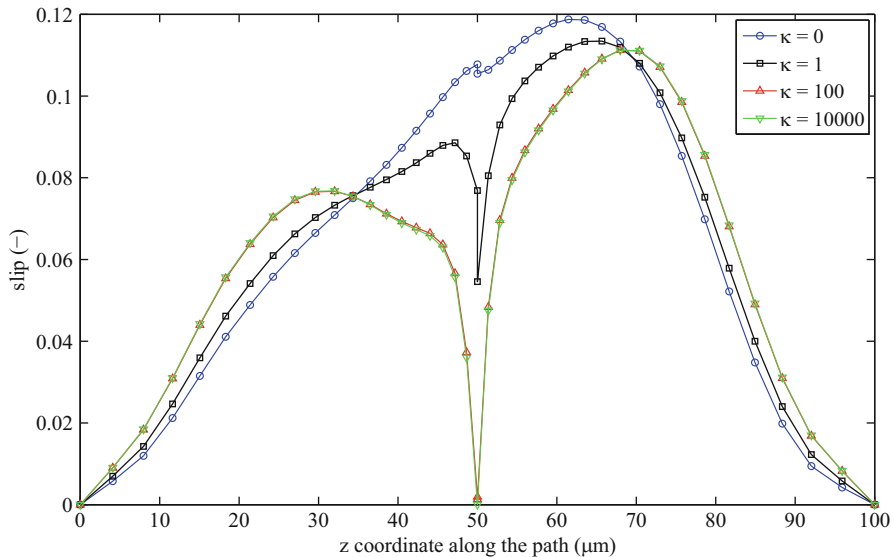


Fig. 7 Slip distribution along the longitudinal path for different κ values

grain boundary resembles to that observed around a hard boundary. The contour plot comparison shown in Fig. 8 indicates that the hard boundary like behavior grows in the vicinity of the whole grain boundary.

To investigate the influence of the “GB strength” on the mechanical response of the system, in Fig. 9, the total reaction force in z-direction is drawn for different κ values. The increase in reaction force and the slope of the force-displacement curves clearly demonstrates the hardening effect of the GB. Due to significant changes in slip distribution, there is a considerable increase in the reaction force. It is also noteworthy to mention that over a certain threshold, the increase in the reaction is

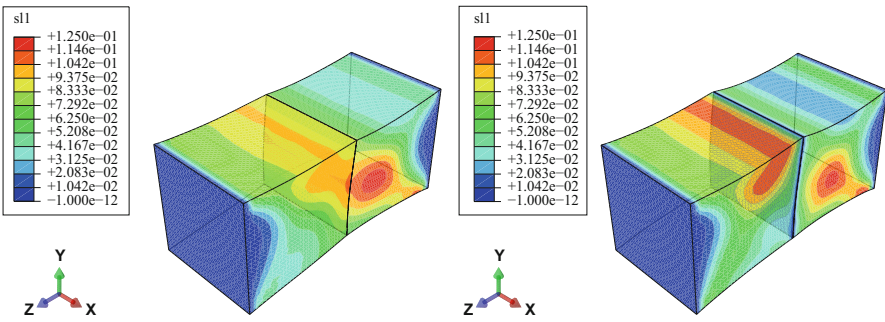


Fig. 8 Slip contours within a crystallographically misaligned bi-crystal for $\kappa = 0$ (left) and $\kappa = 10,000$ (right)

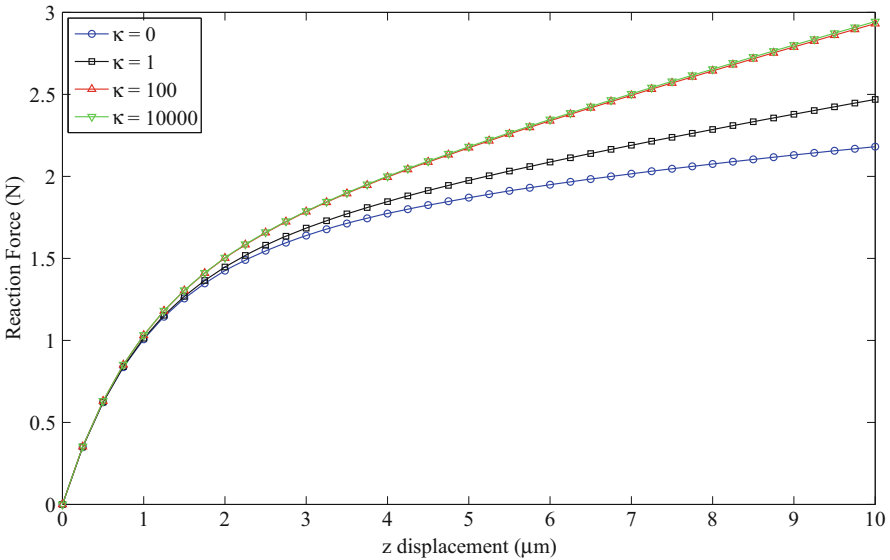


Fig. 9 Total force versus displacement for different κ values when there is a misalignment between the two grains

becoming insignificant which indicates that the grain boundary is sufficiently close to a hard boundary.

Thereafter, the tilted grain boundary with misaligned grains is considered, and the same set of analysis is reiterated. An important and interesting case can be constructed if one of the grains is rotated such that the slip system within that grain becomes parallel to the grain boundary. To set up such a configuration, grain 1 is rotated by 15° counterclockwise. In this case, the slip system in grain 1 is not going to interact with the grain boundary since \vec{s} vector of the slip plane becomes parallel to the GB. Contour plots and the slip profile along the longitudinal path presented in Figs. 10 and 11 are obtained by setting $\kappa = 0$ and $\kappa = 1,000,000$, respectively.

Comparison of the slip distribution for grain 1 for $\kappa = 0$ and $\kappa = 1,000,000$ indicates that the slip distributions are very close in these two cases. In other words, the slip profile in grain 1 that would be obtained without the grain boundary ($\kappa = 0$) is almost identical to the profile obtained when there is almost a hard boundary ($\kappa = 1,000,000$). Such a situation can arise if and only if the crystallographic system of the grain does not interact with the GB. This is in fact the case since the slip system of grain 1 is parallel to the GB. The slight disagreement between the slip profiles of grain 1 for $\kappa = 0$ and $\kappa = 1,000,000$ stems from the fact that the slip distribution in grain 2 is severely modified (for $\kappa = 1,000,000$) which has some consequences on elastic strains and stresses in both grains due to coupled nature of the problem. Due to changes in elastic strains within grain 2, the slip distribution within grain 1 for $\kappa = 0$ and $\kappa = 1,000,000$ is slightly altered although the slip system of this grain and the GB are parallel.

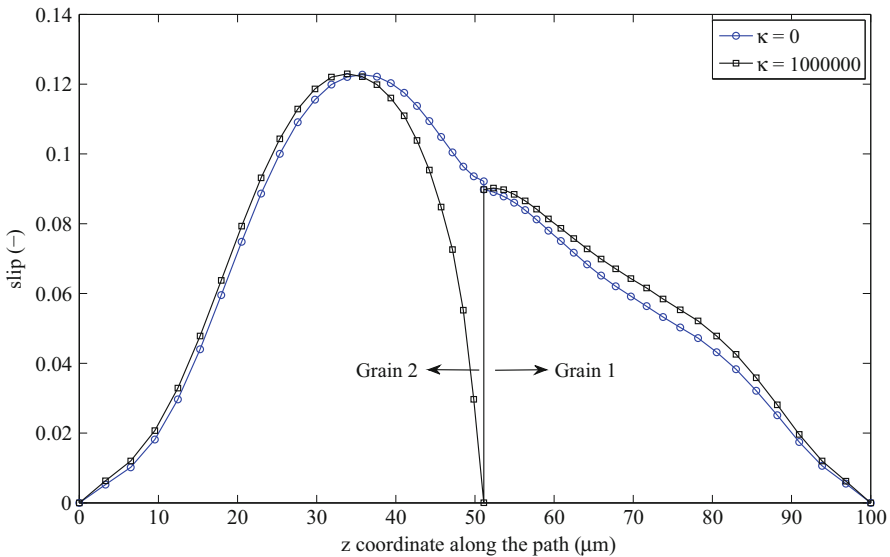


Fig. 10 Slip distribution along the longitudinal axis for $\kappa = 0$ and $\kappa = 1,000,000$ values. Slip distributions within grain 1 are very close to each other for two different κ values

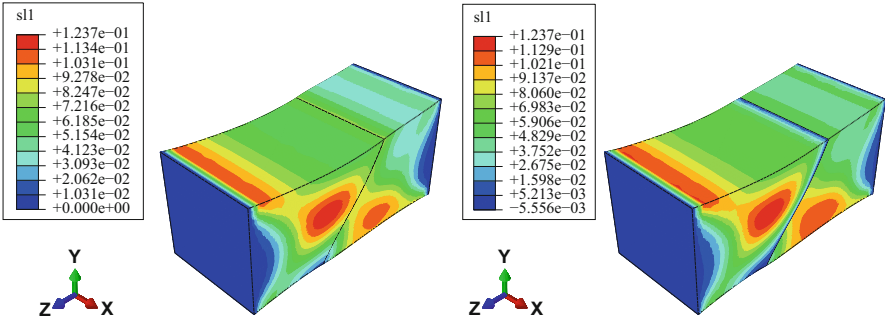


Fig. 11 Slip contours for $\kappa = 0$ (top) and $\kappa = 1,000,000$ (bottom) when the slip system in grain 1 is parallel to the GB

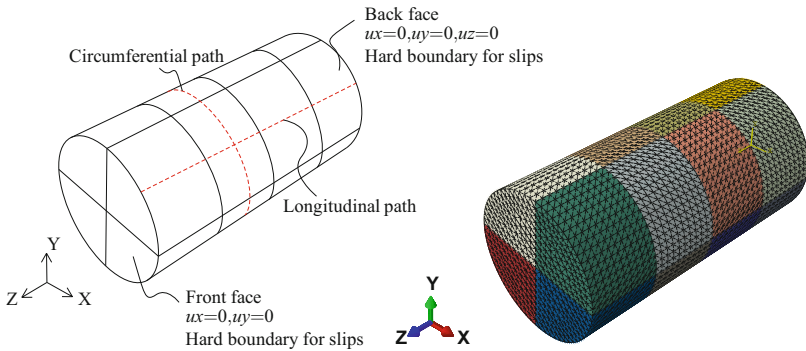


Fig. 12 A cylindrical specimen with 16 grains. The specimen is discretized by 69235 bulk and 74719 interface elements

Cylindrical Specimen with Three Slip System

Although bi-crystal specimens are ideal for investigating grain boundary mechanics, even the smallest engineering components typically consist of multiple grains. Furthermore, a single slip system is far from being realistic. To assess the potential consequences of embedding a grain boundary model within a polycrystalline specimen with multiple active slip systems, a cylindrical specimen consisting of 16 regular grains is considered in this example. The specimen is $12.5 \mu\text{m}$ in radius and $50 \mu\text{m}$ in length and stretched in positive Z-direction by $5 \mu\text{m}$ in 1 s. Please see Fig. 12 for displacement and slip boundary conditions. The parameters of the problem are given in Table 1, and it is assumed that three slip systems (octahedral plane slip systems) are active at any material point as tabulated in the same table. The orientations of the grains are described by Euler angles, and they are generated randomly using the Bunge convention. Therefore the slip systems of the neighboring grains are misaligned in a random manner.

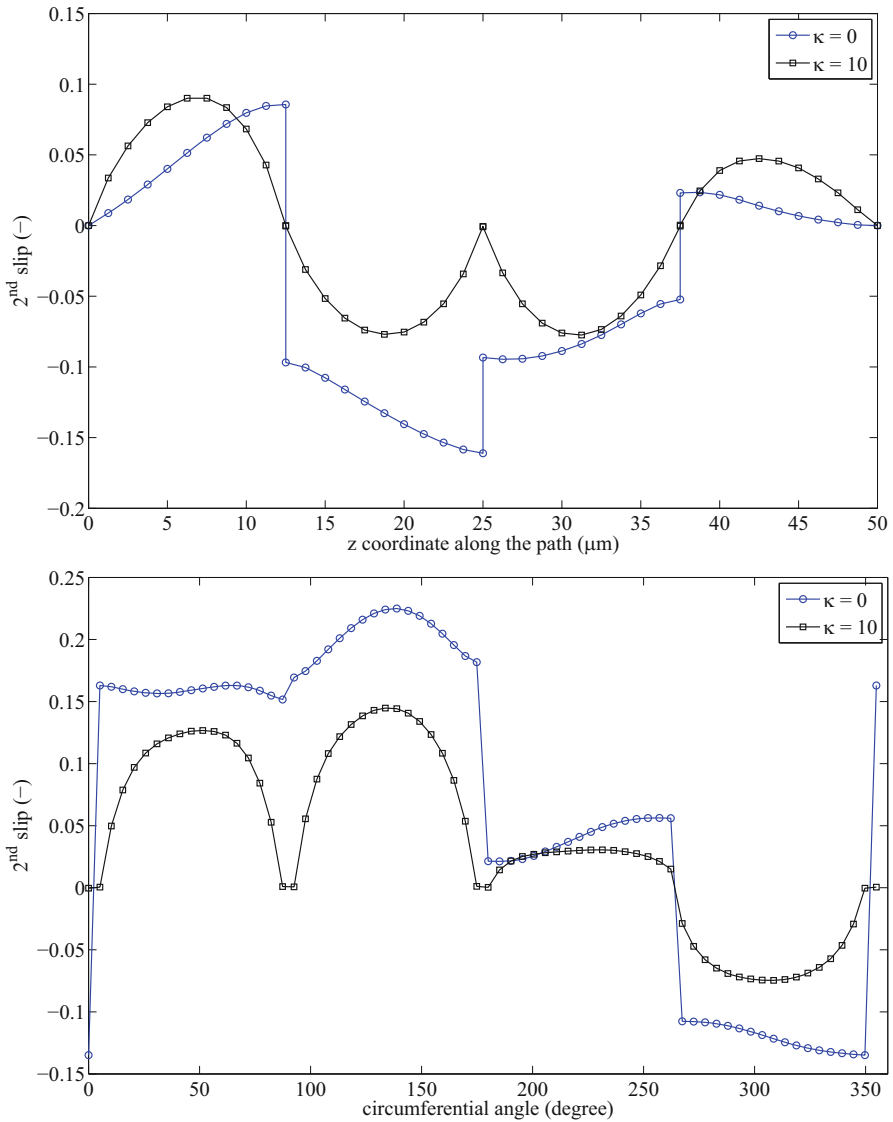


Fig. 13 Slip profile along longitudinal (*top*) and circumferential path (*bottom*) for cylindrical specimen

Two analyses with $\kappa = 0$ and $\kappa = 10$ are conducted, and the corresponding slip distributions (second slip system) along a representative longitudinal and a circumferential path (please see Fig. 12 for the chosen paths) are presented in Fig. 13. Although the same set of Euler angles are used in the two analyses, a stronger grain boundary yields a significant change in slip distribution. The discontinuities

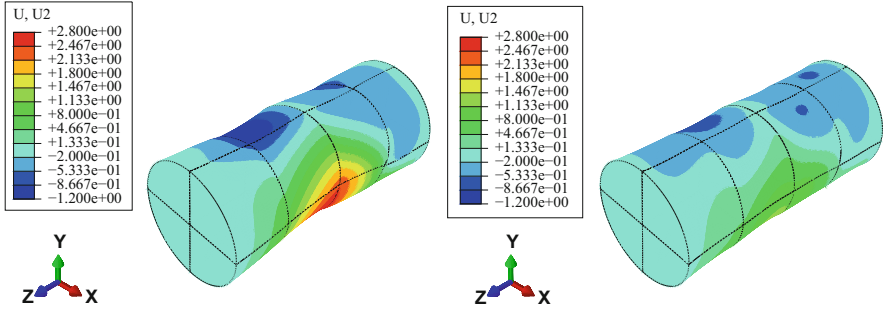


Fig. 14 Y-displacement as obtained by $\kappa = 0$ (left) and $\kappa = 10$ (right). The orientations of the grains in both analyses are kept the same

disappear for both longitudinal and circumferential paths, and a sort of pattern induced by grain boundaries is visible for strong GBs. Most of the grain boundaries behave very similar to hard boundaries although the GB strength is not that high (as compared to the bi-crystal example). Therefore, it seems that not only the strength of the GB but also the misorientation angle between the grains is influential on the slip profiles. Although the second slip system is presented, similar patterns are observed also for slip systems 1 and 3.

So far the discussion has focused on the slip response and the way it is altered as a function of GB strength and misorientations. However, due to coupled nature of the problem, it is natural to expect that GBs would also influence the strain and displacement fields of the specimen. In order to underscore this fact, in Fig. 14, contours of Y-displacement as obtained from the two analyses are presented. The significance of the difference is clearly visible, and in fact this comparison suggests that the GBs might have a strong influence on localization bands in polycrystalline specimens. Although the results of the current analysis are quite limited, the framework has the potential to investigate the influence of GBs on overall behavior to a certain extent through a statistical perspective. Furthermore, by switching on and off the GBs in combination with some specific textures (in terms of grain orientation distributions), the current formulation might be useful to quantify the relative contribution of GBs on mechanical response indicators such as load capacity corresponding to a “macroscopic” strain and/or “macroscopic” strain necessary to reach a certain “microscopic” strain within the polycrystalline solid body.

Conclusion and Outlook

In this chapter, a grain boundary model, which was rigorously derived in Gurtin (2008), is incorporated into a strain gradient crystal plasticity framework in a thermodynamically consistent manner. The description of the grain boundary model

is based on the grain boundary Burgers tensor which takes into account both the effect of the mismatch between the grains and the orientation of the grain boundary.

Macroscopically, the geometric structure of a grain boundary is described by the grain misorientation and the grain boundary normal vector, Wolf and Yip (1992). Therefore the current formulation fully captures the geometric structure of a grain boundary. However, as long as intrinsic energy and energy stemming from interaction of lattice dislocations and grain boundary are concerned, the quadratic energy form used here is not really based on physical arguments. Nonetheless, this form can still be used to investigate the basic properties and limiting cases in a systematic manner. At this stage, it has to be noted that one can construct situations in which grain boundary Burgers tensor is zero although there exist nonzero slips within the grains. A physical explanation of such a situation might correspond to annihilation. However, in that case the model cannot reproduce the hard boundary conditions regardless of the magnitude of κ ; see Gottschalk et al. (2016).

In order to capture the role of grain boundaries more accurately within continuum scale plasticity models, reflecting complicated physics of grain boundaries through proper upscaling strategies is essential. A particular approach in this context makes use of distributed disclination model to describe the initial energy of the symmetric tilt boundaries, Fressengeas et al. (2014). As compared to the results of atomistic models, the predictions of such defect mechanics based models are satisfactory. However, evolution of grain boundary intrinsic energy and incorporation of the effect of lattice dislocation interactions with grain boundaries and the associated energy still need to be addressed. Atomistic studies on grain boundary structures in various material systems are quite numerous, and their results can be exploited to construct computationally feasible multi-scale models. With regard to this, a systematic approach is presented in a series of papers by van Beers et al. (2015a,b) where a logarithmic relation for the initial intrinsic energy with two fitting parameters identified by atomistic simulations is embedded into a strain gradient crystal plasticity model. Furthermore, the interaction with lattice dislocations and grain boundary is quantified in terms of a net defect density measure (consistent with the grain boundary Burgers tensor), and defect redistribution within the grain boundary is addressed through net defect balance equation. Although the model has some phenomenological aspects, some qualitative agreement with small-scale experiments is achieved. Other mechanisms such as grain boundary sliding and opening resulting in the change of the structure of the grain boundary are intimately coupled with defect redistribution. Such changes involving relative sliding and opening call for a proper description of the mechanical response of the grain boundary. The treatment of GBs under finite strains (see, e.g., McBride et al. 2016) and coupling defect redistribution with mechanical cohesive behavior in a thermodynamically consistent manner (see, e.g., Mosler and Scheider 2011) are some open problems to be addressed. Here defect redistribution is primarily concerned with the evolution of the initial structure of the GBs and does not deal with the slip transmission events. It seems that continuum level description of such transmission events requires some special attention on different aspects ranging from the determination of the essential interactions (e.g., dislocation interactions

across a GB, see Stricker et al. (2016) to proper set of transmission criteria; see Bayerschen et al. (2016). Apart from all these potential improvements, it has to be noted that quantitative validation of GB models at continuum scale is a challenge due to intrinsic coupling between defect transport including redistribution and the mechanical integrity of GBs.

References

- R.K. Abu Al-Rub, Interfacial gradient plasticity governs scale-dependent yield strength and strain hardening rates in micro/nano structured metals. *Int. J. Plast.* **24**, 1277–1306 (2008)
- K.E. Aifantis, J.R. Willis, The role of interfaces in enhancing the yield strength of composites and polycrystals. *J. Mech. Phys. Solids* **53**, 1047–1070 (2005)
- K.E. Aifantis, W.A. Soer, J.T. de Hosson, J. Willis, Interfaces within strain gradient plasticity: theory and experiments. *Acta Mater.* **54**, 5077–5085 (2006)
- E. Bayerschen, A.T. McBride, B.D. Reddy, T. Böhlke, Review of slip transmission criteria in experiments and crystal plasticity models. *J. Mater. Sci.* **51**, 2243–2258 (2016)
- C.J. Bayley, W.A.M. Brekelmans, M.G.D. Geers, A comparison of dislocation induced back stress formulations in strain gradient crystal plasticity. *Int. J. Solids Struct.* **43**, 7268–7286 (2006)
- U. Borg, A strain gradient crystal plasticity analysis of grain size effects in polycrystals. *Eur. J. Mech. A. Solids* **26**, 313–324 (2007)
- T. Borg, N.A. Fleck, Strain gradient effects in surface roughening. *Model. Simul. Mater. Sci. Eng.* **15**, S1–S12 (2007)
- M. de Koning, R. Miller, V.V. Bulatov, F.F. Abraham, Modelling grain boundary resistance in intergranular dislocation slip transmission. *Philos. Mag. A* **82**, 2511–2527 (2002)
- M. de Koning, R.J. Kurtz, V.V. Bulatov, C.H. Henager, R.G. Hoagland, W. Cai, M. Nomura, Modeling of dislocation-grain boundary interactions. *J. Nucl. Mater.* **323**, 281–289 (2003)
- M. Ekh, S. Bargmann, M. Grymer, Influence of grain boundary conditions on modeling of size-dependence in polycrystals. *Acta Mech.* **218**, 103–113 (2011)
- N.A. Fleck, J.W. Hutchinson, A formulation of strain gradient plasticity. *J. Mech. Phys. Solids* **49**, 2245–2271 (2001)
- P. Fredriksson, P. Gudmundson, Size-dependent yield strength of thin films. *Int. J. Plast.* **21**, 1834–1854 (2005)
- C. Fressengeas, V. Taupin, L. Capolunga, Continuous modeling of the structure of symmetric tilt boundaries. *Int. J. Solids Struct.* **51**, 1434–1441 (2014)
- M.G.D. Geers, W.A.M. Brekelmans, C.J. Bayley, Second-order crystal plasticity: internal stress effects and cyclic loading. *Model. Simul. Mater. Sci. Eng.* **15**, 133–145 (2007)
- D. Gottschalk, A. McBride, B.D. Reddy, A. Javili, P. Wriggers, C.B. Hirschberger, Computational and theoretical aspects of a grain-boundary model that accounts for grain misorientation and grain-boundary orientation. *Comput. Mater. Sci.* **111**, 443–459 (2016)
- P. Gudmundson, A unified treatment of strain gradient plasticity. *J. Mech. Phys. Solids* **52**, 1379–1406 (2004)
- M.E. Gurtin, On the plasticity of single crystals: free energy, microforces, plastic-strain gradients. *J. Mech. Phys. Solids* **48**, 989–1036 (2000)
- M.E. Gurtin, A gradient theory of single-crystal viscoplasticity that accounts for geometrically necessary dislocations. *J. Mech. Phys. Solids* **50**, 5–32 (2002)
- M.E. Gurtin, A theory of grain boundaries that accounts automatically for grain misorientation and grain-boundary orientation. *J. Mech. Phys. Solids* **56**, 640–662 (2008)
- M.E. Gurtin, L. Anand, S.P. Lele, Gradient single-crystal plasticity with free energy dependent on dislocation densities. *J. Mech. Phys. Solids* **55**, 1853–1878 (2007)
- R. Kumar, F. Szekely, E. Van der Giessen, Modelling dislocation transmission across tilt grain boundaries in 2D. *Comput. Mater. Sci.* **49**, 46–54 (2010)

- G. Lancioni, T. Yalçinkaya, A. Cocks, Energy-based non-local plasticity models for deformation patterning, localization and fracture. *Proc. R. Soc. Lond. A Math. Phys. Eng. Sci.* **471**(2180) (2015a)
- G. Lancioni, G. Zitti, T. Yalcinkaya, Rate-independent deformation patterning in crystal plasticity. *Key Eng. Mater.* **651–653**, 944–949 (2015b)
- T.C. Lee, I.M. Robertson, H.K. Birnbaum, Prediction of slip transfer mechanisms across grain boundaries. *Scr. Metall.* **23**(5), 799–803 (1989)
- Z. Li, C. Hou, M. Huang, C. Ouyang, Strengthening mechanism in micro-polycrystals with penetrable grain boundaries by discrete dislocation dynamics simulation and Hall-Petch effect. *Comput. Mater. Sci.* **46**, 1124–1134 (2009)
- A. Ma, F. Roters, D. Raabe, On the consideration of interactions between dislocations and grain boundaries in crystal plasticity finite element modeling – theory, experiments, and simulations. *Acta Mater.* **54**, 2181–2194 (2006)
- T.J. Massart, T. Pardoen, Strain gradient plasticity analysis of the grain-size-dependent strength and ductility of polycrystals with evolving grain boundary confinement. *Acta Mater.* **58**, 5768–5781 (2010)
- A.T. McBride, D. Gottschalk, B.D. Reddy, P. Wriggers, A. Javili, Computational and theoretical aspects of a grain-boundary model at finite deformations. *Tech. Mech.* **36**, 102–119 (2016)
- D.L. McDowell, Viscoplasticity of heterogeneous metallic materials. *Mater. Sci. Eng. R* **62**, 67–123 (2008)
- J. Mosler, I. Scheider, A thermodynamically and variationally consistent class of damage-type cohesive models. *J. Mech. Phys. Solids* **59**(8), 1647–1668 (2011)
- I. Özdemir, T. Yalçinkaya, Modeling of dislocation-grain boundary interactions in a strain gradient crystal plasticity framework. *Comput. Mech.* **54**, 255–268 (2014)
- Z. Shen, R.H. Wagoner, W.A.T. Clark, Dislocation pile-up and grain boundary interactions in 304 stainless steel. *Scr. Metall.* **20**(6), 921–926 (1986)
- D.E. Spearot, D.L. McDowell, Atomistic modeling of grain boundaries and dislocation processes in metallic polycrystalline materials. *J. Eng. Mater. Tech.* **131**, 041,204 (2009)
- M. Stricker, J. Gagel, S. Schmitt, K. Schulz, D. Weygand, P.B. Gumbsch, On slip transmission and grain boundary yielding. *Meccanica* **51**, 271–278 (2016)
- M.A. Tschopp, D.L. McDowell, Asymmetric tilt grain boundary structure and energy in copper and aluminium. *Philos. Mag.* **87**, 3871–3892 (2007)
- P.R.M. van Beers, G.J. McShane, V.G. Kouznetsova, M.G.D. Geers, Grain boundary interface mechanics in strain gradient crystal plasticity. *J. Mech. Phys. Solids* **61**, 2659–2679 (2013)
- P.R.M. van Beers, V.G. Kouznetsova, M.G.D. Geers, Defect redistribution within continuum grain boundary plasticity model. *J. Mech. Phys. Solids* **83**, 243–262 (2015a)
- P.R.M. van Beers, V.G. Kouznetsova, M.G.D. Geers, M.A. Tschopp, D.L. McDowell, A multiscale model to grain boundary structure and energy: from atomistics to a continuum description. *Acta Mater.* **82**, 513–529 (2015b)
- E. Van der Giessen, A. Needleman, Discrete dislocation plasticity: a simple planar model. *Model. Simul. Mater. Sci. Eng.* **3**, 689–735 (1995)
- D. Wolf, S Yip, *Material Interfaces: Atomic-Level Structure and Properties* (Chapman and Hall, London, 1992)
- T. Yalçinkaya, Multi-scale modeling of microstructure evolution induced anisotropy in metals. *Key Eng. Mater.* **554–557**, 2388–2399 (2013)
- T. Yalçinkaya, W.A.M. Brekelmans, M.G.D. Geers, Non-convex rate dependent strain gradient crystal plasticity and deformation patterning. *Int. J. Solids Struct.* **49**, 2625–2636 (2012)
- T. Yalcinkaya, G. Lancioni, Energy-based modeling of localization and necking in plasticity. *Procedia Mater. Sci.* **3**, 1618–1625 (2014)
- T. Yalcinkaya, W.A.M. Brekelmans, M.G.D. Geers, Deformation patterning driven by rate dependent non-convex strain gradient plasticity. *J. Mech. Phys. Solids* **59**, 1–17 (2011)



Microplane Models for Elasticity and Inelasticity of Engineering Materials

30

Ferhun C. Caner, Valentín de Carlos Blasco,
and Mercè Ginjaume Egido

Contents

Introduction	1066
The Principal Microplane Formulations	1069
The Kinematic No-Split Formulation	1070
The Kinematic Split Formulation	1072
The Static Split Formulation	1074
Microplane Models for Cementitious Materials	1075
Earlier Microplane Models with Limited Predictive Capabilities	1075
The Microplane Model M4	1077
The Microplane Model M7	1078
Nonlocal Microplane Models	1085
A Mesh Size-Independent Microplane Model: The Model M7Auto	1086
Conclusions	1095
References	1095

Abstract

In the traditional approach to the modeling of mechanical behavior of engineering materials, the stress tensor is calculated directly from the prescribed strain tensor either by a closed form tensorial relation as in elasticity or by incremental

F. C. Caner (✉)

School of Industrial Engineering, Institute of Energy Technologies, Universitat Politècnica de Catalunya, Barcelona, Spain

Department of Materials Science and Metallurgical Engineering, Universitat Politècnica de Catalunya, Barcelona, Spain
e-mail: ferhun.caner@upc.edu

V. de Carlos Blasco · M. G. Egido

School of Industrial Engineering, Institute of Energy Technologies, Universitat Politècnica de Catalunya, Barcelona, Spain

analysis as in classical plasticity formulations in which the formulation is developed in terms of tensor invariants and their combinations. However, to model the general three-dimensional constitutive behavior of the so-called geomaterials at arbitrary nonproportional load paths that frequently arise in dynamic loadings, such direct approaches do not yield models with desired accuracy. Instead, microplane approach prescribes the constitutive behavior on planes of various orientations of the material microstructure independently, and the second-order stress tensor is obtained by imposing the equilibrium of second-order stress tensor with the microplane stress vectors. In this work, particular attention is devoted to the milestone microplane models for plain concrete, namely, the model M4 and the model M7. Furthermore, a novel autocalibrating version of the model M7 called the model M7Auto is presented as an alternative to both differential and integral type nonlocal formulations since the model M7Auto does not suffer from the shortcomings of these classical nonlocal approaches. Examples of the performance of the models M7 and M7Auto are shown by simulating well-known benchmark test data like three-point bending size effect test data of plain concrete beams using finite element meshes of the same element size and Nooru-Mohamed test data obtained at different load paths using finite element meshes having different element sizes, respectively.

Keywords

Constitutive model · Microplane model · Crack band model · Concrete · Auto-calibrating Microplane model · Three dimensional finite element analysis

Introduction

The classical approach to the constitutive modeling of engineering materials is to relate the second-order strain tensor to the second-order stress tensor meanwhile satisfying tensorial invariance requirements. Often this approach requires either the use of tensor polynomial or scalar loading functions and inelastic potential functions that vary only with the invariants of the stress or strain tensors. In addition, in the case of pressure-sensitive dilatant quasi-brittle materials such as concretes, rocks, and stiff soils, the principal directions of second-order stress and strain tensors almost never coincide when a general dynamic loading causing significant wave propagation is considered. Thus, the classical approach to constitutive modeling becomes extremely difficult if not impossible in the case of such materials. A multitude of models exist that follow this classical approach, but almost all of them fall short of a complete description of the behavior of quasi-brittle materials at general nonproportional load paths.

In microplane approach, however, the material behavior is prescribed on planes of many different orientations independently, and the response is assembled to yield the second-order stress or strain tensor by arguing equilibrium of stresses that act at different scales in the microstructure of the material. This indirect approach is computationally much more demanding than the classical approach, but at the same

time, it offers many more possibilities for modeling complex material behavior. For example, the so-called kinematic constraint employed for quasi-brittle materials, which implies that the microplane strains on different microplanes are projections of the second-order strain tensor, combined with equivalence of the virtual work of second-order stress and strain tensor over the volume of a unit hemisphere with that of microplane stress and strains over the surface of the same unit hemisphere (called the “stress equilibrium”) automatically yields a microplane model in which the principal directions of the second-order stress and strain tensors almost never coincide. Moreover this microplane model captures the so-called vertex effect automatically (Caner et al. 2002). Similarly, using the so-called static constraint employed for polycrystalline metals, which implies that the microplane stresses on different microplanes are projections of the second-order stress tensor, leads to a multisurface plasticity model after imposing the stress equilibrium. Furthermore, in such a multisurface plasticity model with as many yield surfaces as the number of microplanes, there is no need to invent new, invariant rules for the activation and deactivation of the many yield surfaces for a multitude of possible load paths under dynamic loads; that is automatically determined by the magnitude of the deviatoric part of the projected microplane stress vectors.

For many materials, the microplane can physically be related to the material microstructure. For example, in polycrystalline plasticity, the microplanes represent compact planes over which dislocation motion takes place more easily than on other planes. Although there is a finite number of such planes for many metals and alloys and thus the material response appears to be likely anisotropic, many different grains oriented in many different directions actually lead to isotropic behavior in the structural scale with dimensions on the order of 0.01 mm or larger. In the case of concretes, there are no dislocations; however, as shown in Fig. 1b, the contact surfaces between coarse and fine aggregates and cement paste within the material mesostructure are where the inelastic phenomena take place, and thus microplanes represent a collection of all such contact planes oriented in all possible directions creating again initially isotropic response in the structural scale with dimensions on the order of 5 mm or larger. In biological soft tissue, for example, the material is made of a matrix of elastin fibers forming a network reinforced with collagen fibers with a mean direction oriented close to the maximum stress direction in the tissue. The distribution of collagen fiber directions can easily be projected on to the microplanes oriented in all possible directions which results in anisotropic behavior. The elastin fiber network is almost completely isotropic and can readily be represented by these microplanes.

The focus of this study is microplane models for plain concrete, arguably one of the most challenging engineering materials to model mathematically. Only the most important milestones in the development of microplane models for concrete are dealt with. In particular, the model M4 (Bažant et al. 2000; Caner and Bažant 2000), the model M7 (Caner and Bažant 2013a,b), and a novel nonlocal extension of the model M7 called the model M7Auto are studied in this work. The models M4 and M7 are continuum models developed to predict fracture, damage, and plasticity in plain concrete using the finite element method in which the damage

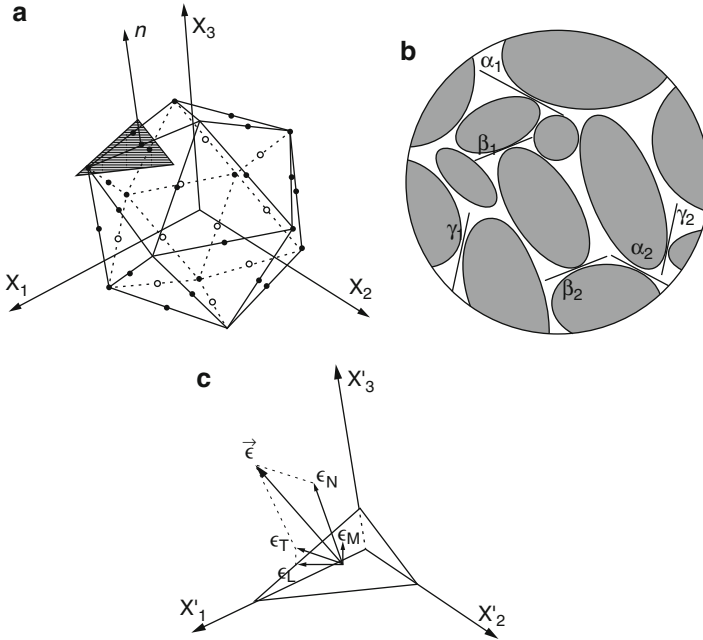


Fig. 1 (a) The Gaussian quadrature points representing the microplanes over a unit sphere, (b) the weak planes shown in the mesostructure of concrete, (c) the microplane strain vectors shown on a microplane

and fracture localize to a one element wide band. Such models are also called crack band models. In these models the response depends on the size and shape of the elements in the finite element mesh employed. In crack band models, the calibration of the model is carried out for a fixed element size, and in the analyses a finite element mesh of the same size elements with an aspect ratio of approximately 1 must be employed to obtain accurate predictions of the concrete response to an arbitrary three-dimensional stress state. By contrast, the classical approach to remove the mesh size dependence of crack band models is in the form of either (1) an integral type approach or (2) a differential type approach in both of which a characteristic material size acts as a localization limiter. Both of these approaches suffer from significant issues such as unobjective treatment of the Gauss points near the boundary of the structure, prediction of cracks initiating ahead of the crack tip, errors introduced due to shielding of Gauss points by a partially open cohesive crack, physically difficult to interpret higher-order stresses, and associated boundary conditions. In this study, the model M7Auto is described to overcome all these difficulties simply by taking into account the element size, shape, and type during the calibration process to extract the dependence of the microplane stress-strain boundaries on the finite element size so as to dissipate a constant fracture energy at quasistatic loading rates. It is shown that to successfully employ the model

M7Auto in the finite element analysis of concrete structures, it is necessary that (1) the size of the elements in the mesh does not exceed the range of autocalibration of the model, (2) the elements have an aspect ratio close to 1 for the simulation of general three-dimensional crack propagation, and (3) the element type should be the same in the mesh. The predictive capabilities of the models M7 and M7Auto are demonstrated by simulating some well-known benchmark test data.

Other than microplane models for cementitious materials, there are a number of microplane models formulated for polymer foams (Brocca et al. 2001), polycrystalline metals (Brocca and Bažant 2000), fiber-reinforced polymer composites (Kirane et al. 2016; Caner et al. 2011; Cusatis et al. 2008), rocks (Bažant and Zi 2003; Chen and Bažant 2014), soils (Prat and Bažant 1991; Chang and Sture 2006), shape memory alloys (Brocca et al. 2002), and biological soft tissue (Caner et al. 2007). Details of these formulations are excluded from the present study.

The Principal Microplane Formulations

Essentially the formulation for the constitutive response of a material involves determination of the stress tensor given the strain tensor and, in the inelastic range, the history variables that describe the earlier state of the material. In general, this relation must be a direct relation as in the case of elastic behavior of materials. Often an incremental approach must be employed in the inelastic range due to several mathematical and physical reasons. A typical example is the classical J_2 plasticity formulation that separates the plastic part from the strain tensor and determines the current stress tensor that corresponds to the current strain tensor. The microplane approach is fundamentally different in that the response quantities are first determined on the microplanes and integrated over all possible microplanes oriented in different directions to yield the familiar second-order tensor form. Due to the relationship between the material microstructure and the mechanical response, the application of the microplane approach to the constitutive modeling of materials with very different mechanical behaviors is of interest. Consequently, several formulations of microplane models proposed for several different engineering materials exist; among them are the so-called geomaterials that include cement-based materials, rocks, and soils. These materials exhibit pressure sensitivity, dilatancy, and softening due to fracturing. To capture this strain-softening behavior, the so-called kinematic constraint must be applied instead of the so-called static constraint. In the kinematically constrained microplane formulations, both the strain tensor and its increment are projected over all possible microplanes as shown in Fig. 1c, and the stress vector acting on each of these microplanes is evaluated using the constitutive law prescribed on the microplane level. Finally, the stress tensor is calculated by integrating the stress vector over all possible microplanes as shown in Fig. 1a. By contrast, in the statically constrained microplane formulations used to describe polycrystalline plasticity of metals and their alloys, both the stress tensor and its increment are projected over all possible microplanes, and the

corresponding strain vectors are calculated using the constitutive law prescribed at the microplane level. Finally, the second-order tensor form of the strain tensor is obtained by integrating the strain vector over all possible microplanes. Thus, the microplane approach is an indirect one that employs simpler constitutive relations between stress and strain vectors at the microplane level to model complex material behavior. The kinematically constrained formulations can further be divided into the formulations with the so-called volumetric-deviatoric split and those without the volumetric-deviatoric split. The split in these formulations refers to the separation of the strain tensor into volumetric and deviatoric parts.

In the remainder of this section, the most commonly used formulations, namely, the kinematic no-split, the kinematic split, and the static split formulations, are discussed.

The Kinematic No-Split Formulation

Given the strain tensor, the first step is to project it onto the microplanes defined by discretizing a unit hemisphere at special locations for optimal integration (e.g., using Gaussian quadrature) over the surface of the hemisphere:

$$\epsilon_N = N_{ij} \epsilon_{ij} \quad (1)$$

where $N_{ij} = n_i n_j$ and the normal vector \mathbf{n} is the normal to the microplane defined for each of the microplanes involved in the model.

To treat the microplane shear strain vector, it is convenient to define the directions \mathbf{m} and \mathbf{l} so that the base vectors in the directions \mathbf{n} , \mathbf{m} , and \mathbf{l} form an orthogonal rectangular coordinate system on the microplane. These directions should be defined randomly so that no directional bias is created across the microplanes that form the unit hemisphere. Using these vectors, we can define the second-order tensors:

$$\begin{aligned} M_{ij} &= \frac{1}{2} (m_i n_j + m_j n_i) \\ L_{ij} &= \frac{1}{2} (l_i n_j + l_j n_i) \end{aligned} \quad (2)$$

Using the tensors \mathbf{M} and \mathbf{L} , the components of shear strain vector along the directions \mathbf{m} and \mathbf{l} can be evaluated as

$$\begin{aligned} \epsilon_M &= M_{ij} \epsilon_{ij} \\ \epsilon_L &= L_{ij} \epsilon_{ij} \end{aligned} \quad (3)$$

The stress quantities that correspond to these microplane strain quantities are evaluated using the microplane constitutive laws:

$$\begin{aligned}
\sigma_N &= \mathcal{F}_N(\epsilon_N, \dots) \\
\tau_M &= \mathcal{F}_\tau(\epsilon_N, \epsilon_M, \dots) \\
\tau_L &= \mathcal{F}_\tau(\epsilon_N, \epsilon_L, \dots)
\end{aligned} \tag{4}$$

where the functions \mathcal{F}_N and \mathcal{F}_τ denote the microplane normal and shear constitutive laws.

Finally to calculate the stress tensor, the following stress equilibrium equation may be employed:

$$\begin{aligned}
\int_V \sigma_{ij} \delta \epsilon_{ij} dV &= \int_S [\sigma_N \delta \epsilon_N + \tau_M \delta \epsilon_M + \tau_L \delta \epsilon_L] dS \tag{5} \\
\Rightarrow \sigma_{ij} \delta \epsilon_{ij} \int_V dV &= \int_S [\sigma_N N_{ij} \delta \epsilon_{ij} + \tau_M M_{ij} \delta \epsilon_{ij} + \tau_L L_{ij} \delta \epsilon_{ij}] dS \\
\Rightarrow \left[\sigma_{ij} \frac{2\pi}{3} - \int_S [\sigma_N N_{ij} + \tau_M M_{ij} + \tau_L L_{ij}] dS \right] \delta \epsilon_{ij} &= 0 \quad \forall \delta \epsilon_{ij} \neq 0 \\
\Rightarrow \sigma_{ij} &= \frac{3}{2\pi} \int_S [\sigma_N N_{ij} + \tau_M M_{ij} + \tau_L L_{ij}] dS \tag{6}
\end{aligned}$$

Eq. (5) means that the virtual work of stress and strain tensor over the volume of the unit hemisphere is equal to the virtual work of the microplane stresses and strains over the surface of the unit hemisphere.

As a useful exercise, the small strain linearly elastic behavior may be obtained using this formulation as follows. First, the microplane stress and strains are assumed to be related as

$$\begin{aligned}
\sigma_N &= E_N \epsilon_N \\
\tau_M &= E_\tau \epsilon_M \\
\tau_L &= E_\tau \epsilon_L
\end{aligned} \tag{7}$$

Substituting Eqs. (7), (1) and (3) into Eq. (6) and factoring out the strain tensor ϵ yield the stiffness tensor as obtained by the microplane no-split formulation. This tensor may be compared to the classical isotropic elastic stiffness tensor, and for the two tensors to be identical, the microplane elastic moduli turn out to be

$$\begin{aligned}
E_N &= \frac{E}{1 - 2\nu} \\
E_\tau &= \frac{E}{1 - 2\nu} \frac{1 - 4\nu}{1 + \nu}
\end{aligned} \tag{8}$$

According to the above relations, the thermodynamically admissible range of the coefficient of the Poisson is limited to $-1 < \nu < 1/4$. Thus for a material for which the coefficient of Poisson lies in the range $[1/4, 1/2)$, the model cannot be used to predict the behavior of that material. However, for many quasi-brittle materials such

as concretes, fiber-reinforced concretes, and many other ceramics, the coefficient of Poisson lies in the range $[-1, 1/4]$, and this formulation can be used. For example, recently this formulation has been successfully extended to model the complete range of the mechanical behavior of normal strength and high strength concretes as well as fiber-reinforced concretes.

The Kinematic Split Formulation

In this formulation, the microplane normal strain is decomposed into volumetric and deviatoric parts:

$$\begin{aligned} \epsilon_N &= \epsilon_V + \epsilon_D \\ \sigma_N &= \sigma_V + \sigma_D \end{aligned} \tag{9}$$

where

$$\epsilon_V = \epsilon_{kk}/3 \tag{10}$$

and thus

$$\epsilon_D = \left(N_{ij} - \frac{\delta_{ij}}{3} \right) \epsilon_{ij} \tag{11}$$

The shear strains remain as given in Eq. (3). However, in this formulation Eqs. (4) should be rewritten as

$$\begin{aligned} \sigma_D &= \mathcal{F}_D (\epsilon_D, \dots) \\ \sigma_V &= \mathcal{F}_V (\epsilon_V, \dots) \\ \tau_M &= \mathcal{F}_\tau (\epsilon_D, \epsilon_V, \epsilon_M, \dots) \\ \tau_L &= \mathcal{F}_\tau (\epsilon_D, \epsilon_V, \epsilon_L, \dots) \end{aligned} \tag{12}$$

Furthermore, Eq. (6) must also be rewritten as

$$\begin{aligned} &\int_V \sigma_{ij} \delta \epsilon_{ij} dV = \int_S [\sigma_V \delta \epsilon_V + \sigma_D \delta \epsilon_D + \tau_M \delta \epsilon_M + \tau_L \delta \epsilon_L] dS \\ \Rightarrow &\sigma_{ij} \delta \epsilon_{ij} \int_V dV = \int_S \left[\sigma_V \frac{\delta_{ij}}{3} \delta \epsilon_{ij} + \sigma_D \left(N_{ij} - \frac{\delta_{ij}}{3} \right) \delta \epsilon_{ij} + \tau_M M_{ij} \delta \epsilon_{ij} + \tau_L L_{ij} \delta \epsilon_{ij} \right] dS \\ \Rightarrow &\left[\sigma_{ij} \frac{2\pi}{3} - \sigma_V \delta_{ij} \frac{2\pi}{3} - \int_S \left[\sigma_D \left(N_{ij} - \frac{\delta_{ij}}{3} \right) + \tau_M M_{ij} + \tau_L L_{ij} \right] dS \right] \delta \epsilon_{ij} = 0 \quad \forall \delta \epsilon_{ij} \neq 0 \\ \Rightarrow &\sigma_{ij} = \sigma_V \delta_{ij} + \frac{3}{2\pi} \int_S \left[\sigma_D \left(N_{ij} - \frac{\delta_{ij}}{3} \right) + \tau_M M_{ij} + \tau_L L_{ij} \right] dS \end{aligned} \tag{13}$$

because $\int_S N_{ij} dS = \delta_{ij} 2\pi/3$ and $\int_S dS = 2\pi$. Instead of Eq. (13) the microplane split formulations prior to the microplane model M4 used simply

$$\sigma_{ij} = \sigma_V \delta_{ij} + \frac{3}{2\pi} \int_S [\sigma_D N_{ij} + \tau_M M_{ij} + \tau_L L_{ij}] dS \quad (14)$$

which violates the work conjugacy of volumetric stress. This can be conveniently shown by simply setting $i = j = k$ in Eq. (14) which yields

$$\sigma_{kk} = 3\sigma_V + \frac{3}{2\pi} \int_S \sigma_D dS \quad (15)$$

in which generally $\int_S \sigma_D dS \neq 0$ except in the elastic range. Thus, the pressure depends on deviatoric stress when Eq. (14) is employed in the calculation of the second-order stress tensor.

Using this formulation the isotropic linearly elastic behavior may be simulated as follows: Let the relationships between the microplane stress and strains be given by

$$\begin{aligned} \sigma_V &= E_V \epsilon_V \\ \sigma_D &= E_\tau \epsilon_D \\ \tau_M &= E_\tau \epsilon_M \\ \tau_L &= E_\tau \epsilon_L \end{aligned} \quad (16)$$

Substituting Eqs. (16), (3), (10), and (11) into Eq. (13) and again factoring out the strain tensor ϵ result in the stiffness tensor as obtained by the microplane split formulation. This tensor may again be compared to the classical elastic stiffness tensor for an isotropic material, and for the two tensors to be identical, the microplane elastic moduli turn out to be

$$\begin{aligned} E_V &= \frac{E}{1 - 2\nu} \\ E_D &= E_\tau = \frac{E}{1 + \nu} \end{aligned} \quad (17)$$

The above equations clearly show that in this formulation, the full thermodynamically admissible range $[-1, 1/2]$ of the coefficient of Poisson can be represented. Several microplane models for concrete and fiber-reinforced concrete have been developed by different authors using this formulation. Due to the volumetric-deviatoric split of the normal strain both in the elastic and in the inelastic ranges, this formulation results in similar material behavior both in tension and in compression. On the other hand, for quasi-brittle materials like concrete, such similarity in the

predicted response is not helpful because the behavior in tension is completely different than behavior in compression for such materials.

The Static Split Formulation

The static split formulation is obtained by projecting the stress tensor over the microplanes:

$$\begin{aligned}\sigma_N &= N_{ij} \sigma_{ij} \\ \tau_M &= M_{ij} \sigma_{ij} \\ \tau_L &= L_{ij} \sigma_{ij}\end{aligned}\quad (18)$$

In this formulation, typically the stress tensor being projected is the trial stress tensor as in the classical plasticity theory. Furthermore, the projected normal stress is divided into volumetric and deviatoric parts:

$$\sigma_N = \sigma_V + \sigma_D \quad (19)$$

where $\sigma_V = \delta_{ij} \sigma_{ij} / 3$ and $\sigma_D = (N_{ij} - \delta_{ij} / 3) \sigma_{ij}$. The projected and split microplane stresses are employed to determine the microplane inelastic strains using the constitutive laws \mathcal{H}_D , \mathcal{H}_V , and \mathcal{H}_τ defined at the microplane level:

$$\begin{aligned}\epsilon_D^{in} &= \mathcal{H}_D(\sigma_D, \sigma_V, \dots) \\ \epsilon_V^{in} &= \mathcal{H}_V(\sigma_V, \dots) \\ \epsilon_M^{in} &= \mathcal{H}_\tau(\sigma_D, \sigma_V, \sigma_M, \dots) \\ \epsilon_L^{in} &= \mathcal{H}_\tau(\sigma_D, \sigma_V, \sigma_L, \dots)\end{aligned}\quad (20)$$

Finally to calculate the inelastic strain tensor, the following stress equilibrium equation may be employed:

$$\begin{aligned}\int_V \epsilon_{ij}^{in} \delta \sigma_{ij} dV &= \int_S [\epsilon_V^{in} \delta \sigma_V + \epsilon_D^{in} \delta \sigma_D + \epsilon_M^{in} \delta \tau_M + \epsilon_L^{in} \delta \tau_L] dS \quad (21) \\ \Rightarrow \epsilon_{ij}^{in} \delta \sigma_{ij} \int_V dV &= \int_S \left[\epsilon_V^{in} \frac{\delta_{ij}}{3} \delta \sigma_{ij} + \epsilon_D^{in} \left(N_{ij} - \frac{\delta_{ij}}{3} \right) \delta \sigma_{ij} + \epsilon_M^{in} M_{ij} \delta \sigma_{ij} + \epsilon_L^{in} L_{ij} \delta \sigma_{ij} \right] dS \\ \Rightarrow \left[\epsilon_{ij}^{in} \frac{2\pi}{3} - \epsilon_V^{in} \delta_{ij} \frac{2\pi}{3} - \int_S \left[\epsilon_D^{in} \left(N_{ij} - \frac{\delta_{ij}}{3} \right) + \epsilon_M^{in} M_{ij} + \epsilon_L^{in} L_{ij} \right] dS \right] \delta \sigma_{ij} &= 0 \quad \forall \delta \sigma_{ij} \neq 0 \\ \Rightarrow \epsilon_{ij}^{in} &= \epsilon_V^{in} \delta_{ij} + \frac{3}{2\pi} \int_S \left[\epsilon_D^{in} \left(N_{ij} - \frac{\delta_{ij}}{3} \right) + \epsilon_M^{in} M_{ij} + \epsilon_L^{in} L_{ij} \right] dS \quad (22)\end{aligned}$$

Eq.(22) determines the inelastic part of the total strain tensor which readily facilitates the calculation of the elastic part of the strain tensor which in turn leads

to the actual stress tensor. For example, assuming a series coupling of elastic and plastic zones in an elastoplastic material, once the inelastic strain tensor is calculated using Eq. (22), the elastic part of the total strain tensor can easily be determined which readily facilitates the calculation of the actual stress tensor.

The microplane isotropic linearly elastic flexibility tensor can be obtained by treating the inelastic strains in Eq. (22) as elastic strains and substituting Eq. (18) and

$$\begin{aligned}\epsilon_V &= \frac{1}{E_V} \sigma_V \\ \epsilon_D &= \frac{1}{E_\tau} \sigma_D \\ \epsilon_M &= \frac{1}{E_\tau} \sigma_M \\ \epsilon_L &= \frac{1}{E_\tau} \sigma_L\end{aligned}\tag{23}$$

into Eq. (22) and factoring out the stress tensor. By comparing this microplane flexibility tensor with the standard isotropic linearly elastic flexibility tensor, the same relations given in Eq. (17) are obtained.

Microplane Models for Cementitious Materials

Several authors have proposed microplane models for constitutive behavior of cement-based materials since the microplane approach with a formulation applicable to tensile fracturing of cement-based materials was first outlined in 1984 (Bažant and Gambarova 1984; Bažant and Oh 1985). In the remainder of this section, the most important formulations of microplane approach for cementitious materials are discussed.

Earlier Microplane Models with Limited Predictive Capabilities

The first microplane no-split-type formulation presented in Bažant and Gambarova (1984) and Bažant and Oh (1985) is based on the hypotheses that the microplane normal strains are the projections of the second-order strain tensor as shown in Eq. (1) and that the microplane normal stress on any given microplane is a function of the microplane normal strain on that microplane as shown in the first of Eq. (4). The shear resistance on the microplanes is neglected. The microplane system then is coupled to an elastic solid in series so as to yield the desired Poisson's ratio for concrete. To numerically calculate the integral in Eq. (6) (where τ_M and τ_L are taken as zero in this formulation) as efficiently as possible, the optimum Gaussian integration over the surface of a sphere is

employed. Using this model, the experimentally observed dilatancy in concrete when cracks are sheared is successfully simulated. In addition, this formulation is shown to be capable of simulating the tensile fracture of concrete. On the other hand, the compression failure of concrete cannot be correctly predicted using this formulation.

In the microplane split-type formulation presented in Bažant and Prat (1988a) and verified in Bažant and Prat (1988b), the microplane shear resistance is taken into account (as in Eqs. (3) with (12)), and the invariants of the stress and strain tensors are introduced in the formulation so as to be able to model the general three-dimensional pressure-sensitive dilatant behavior of concrete in compression. The volumetric-deviatoric split of the microplane normal strain and stress (as in Eqs. (9) and (19)) is introduced leading to the simulation of the complete thermodynamically admissible range of Poisson's ratio. With this formulation, the microplane formulation is able to predict a considerably wide range of inelastic behavior of concrete, including tensile softening, dilatancy under shear, dilatancy and friction in compression, compression softening, ductile-brittle transition in compression, and hydrostatic compression behaviors. However, this formulation cannot accurately predict the concrete response to load cycles in compression and in tension. Furthermore, this formulation suffers from excessively long softening tails in the model response.

The microplane split-type formulation presented in Ožbolt and Bažant (1992) refines and extends the model presented in Bažant and Prat (1988a) to cyclic loading and the strain rate effect in concrete. The rate effect is modeled as in the Maxwell series coupling model of a nonlinear spring and a dashpot which ignores the real physical mechanisms behind this phenomenon. Although some success is achieved in simulating the load cycles in compression, the load cycles in tension could not be correctly simulated. In particular, this formulation fails to predict damage in tension manifested by unloading to origin with little observed plasticity. In this formulation the microplane damage variables are treated as nonlocal variables as described in Bažant and Ožbolt (1990) and in Bažant and Pijaudiercabot (1988).

The studies Bažant et al. (1996a,b) present and verify yet another split-type formulation that introduces the so-called stress-strain boundaries. These boundaries should be considered as yield limits that vary with microplane strains and other history variables defined on the microplane. The basic idea is to describe the monotonic loading in inelastic range using these boundaries and prescribe the loading, unloading, and reloading responses between the boundaries as elastic. It is verified against different test data from different researchers as well as several test data obtained using specimens from the same concrete at different proportional and nonproportional load paths by the same researchers. Remarkably all these data are simulated relatively well using this formulation. However, in this model the prediction of concrete response to load cycles in tension still remains unsolved. Furthermore, the stiffness degradation in tensile softening cannot be simulated. Also, in tensile softening predictions, the lateral strains are observed to be excessively large which cannot be physically justified.

The Microplane Model M4

The microplane model M4 deserves more attention as it is currently the most widely used microplane model for concrete even though it is *not* the most capable microplane model currently available for concrete. It has been implemented in many commercial and free finite element analysis codes including ATENA, OOFEM, MARS, DIANA, SBETA, and EPIC with an ever-growing user community because it satisfies the needs of many engineers and researchers by providing consistently outstanding prediction of concrete behavior at most commonly used load paths.

One important feature of the model M4 is, according to the model M4 formulation presented in Bažant et al. (2000) and verified in Caner and Bažant (2000), that the work conjugacy of volumetric stress is satisfied leading to the stress equilibrium equation given by Eq. (13). All split-type formulations prior to the model M4 suffer from a lack of work conjugacy of volumetric stress, the correction of which is the main motivation behind the model M4. When Eq. (14) is used in the calculation of the second-order stress tensor instead of Eq. (13), a volumetric-deviatoric coupling occurs automatically in the inelastic range. This, on the other hand, is *not* an error per se because the experimentally observed dilatant behavior of concrete under shear in fact implies the existence of such a coupling. However, this coupling must be introduced explicitly in the boundaries of the model so as to control the model behavior to fit the test data more accurately. Consequently, the model M4 formulation employs Eq. (13) that satisfies the work conjugacy of the volumetric stress.

Another feature of the model M4 is that the microplane boundaries that control softening are enhanced with plateaus so that the sharp peaks observed in the softening predictions by the earlier formulations become rounded in the softening response of the model. In the earlier formulations, the inelastic strains localize to a few microplanes which create a sharp peak which is observed in the experiments as a round one. In the formulation of the model M4, the plateaus incorporated into the microplane boundaries that control the softening response allow more microplanes to reach the yield limit, and thus a rounded peak results.

Yet another major feature embedded in the model M4 is the introduction of a completely new nonlinear friction boundary. Earlier formulations use a linear dependence of the frictional stress on the normal stress which spuriously strengthens the concrete at large normal stresses. The newly introduced boundary starts as nearly linear at small normal stresses but quickly reaches a completely horizontal asymptote at large normal stresses. Thus, at large confining pressures, a maximum frictional resistance at the microplanes is reached. This feature allows the prediction of both the entry and exit crater shapes in the dynamic projectile penetration simulations of concrete walls of different thicknesses which involve only moderately large microplane normal stresses (Bažant et al. 2000). With the microplane formulations prior to the model M4, these crater shapes are not predicted correctly at all.

Furthermore, the unloading and reloading laws which remain elastic in the formulation presented in Bažant et al. (1996b) are redefined to fit the experimental data better in compression load cycles as well as in unloading at high confining pressures.

In tension, a crack-closing boundary is proposed which allows the response to return to origin when unloading occurs. However, the resulting load cycle loops are too large compared to the experimentally observed loops. Furthermore, the stiffness degradation in tension is not captured at all in this formulation. Yet another problem that plagues the model M4 as well as the earlier split formulations is that in the tensile softening response, excessive lateral contractions which cannot be physically justified are predicted.

The model M4 is extended to strain rate effects and finite strain as well, but the important feature of this extension given in Bažant et al. (2000) is that the fracturing strain rate effects are based on the theory of Arrhenius-type fracture process with an activation energy. Although this theory is empirical, it can describe an incredibly wide range of physical phenomena including elastic deformations as well as diffusion phenomena in the material microstructure. The finite strain theory presented and applied to the microplane model M4 in Bažant et al. (2000) proposes as the finite strain tensor the Green's Lagrangian strain tensor and as the stress tensor the back-rotated Cauchy stress tensor which are *not* work conjugate. However, the error that results from the lack of work conjugacy is predicted to be small.

The Microplane Model M7

The microplane model M7 presented in Caner and Bažant (2013a) and verified in Caner and Bažant (2013b) is currently the most capable of the microplane models for concrete. Although it has a kinematic no-split-type formulation in the elastic range, in the inelastic range, the compressive stress-strain boundaries employ the volumetric-deviatoric split of microplane stress and strains, while the tensile boundary still employs the total microplane normal strain, and thus its formulation is of mixed type. The formulation of this model resolves the three aforementioned issues that have been persistent for several decades, namely, (1) the excessive lateral contraction during softening in tension is prevented (now the lateral strains return to origin as the tensile softening progresses), (2) the stiffness degradation during tensile softening is correctly captured, and (3) the tension-compression load cycles are correctly represented. In the rest of the load paths, the response predicted by the model M7 is either slightly better or about the same as that of the model M4. Thus the model M7 appears to be a significant improvement over its predecessors. The price to pay to achieve it is that the Poisson's ratio must stay in the interval $(-1, 1/4)$, which typically is the case for all kinematic no-split microplane formulations that do not ignore the shear resistance on the microplanes. This situation, however, could be remedied by coupling an elastic solid in series to the microplane solid as shown in Caner and Bažant (2013a) which allows the representation of the full thermodynamically admissible range of Poisson's ratio by this series coupling model as a whole.

The stress equilibrium equation employed in this formulation is the same as in Eq. (6). In the elastic range, Eqs. (7) and (8) hold true. In the inelastic range, the microplane stress and strains are separated into their volumetric and deviatoric parts as in Eqs. (9) and used in the microplane constitutive laws given by

$$\begin{aligned}
\mathcal{F}_D(\epsilon_D, \epsilon_V) &= -\frac{Ek_1\beta_3}{1 + ((-\epsilon_D)/(k_1\beta_2))^2} \\
\mathcal{F}_V(\epsilon_V, \epsilon_I, \epsilon_{III}) &= -Ek_1k_3 \exp\left(-\frac{\epsilon_V}{k_1\alpha(\epsilon_I, \epsilon_{III})}\right) \\
\mathcal{F}_N^-(\epsilon_D, \epsilon_V, \epsilon_I, \epsilon_{III}) &= \mathcal{F}_D(\epsilon_D, \epsilon_V) + \mathcal{F}_V(\epsilon_V, \epsilon_I, \epsilon_{III}) \\
\mathcal{F}_N^+(\epsilon_N, \sigma_V) &= Ek_1\beta_1 \exp\left(-\langle\epsilon_N - \beta_1c_2k_1\rangle / (c_4\epsilon_e^0 + k_1c_3)\right) \\
\mathcal{F}_\tau(\sigma_N, \epsilon_V) &= \frac{c_{10}\langle\hat{\sigma}_N^0(\epsilon_V) - \langle-\sigma_N\rangle\rangle}{1 + c_{10}\langle\hat{\sigma}_N^0(\epsilon_V) - \langle-\sigma_N\rangle\rangle / (E_\tau k_1 k_2)}
\end{aligned} \tag{24}$$

where $\epsilon_e^0 = \sigma_V^0/E_{N0}$, $\beta_1 = -c_1 + c_{17} \exp[-(c_{19}\langle-\sigma_V^0 - c_{18}\rangle)/E_{N0}]$, $\beta_2 = c_5 \exp(f'_{c0}/E_0 - f'_c/E) \times \tanh(c_9\langle-\epsilon_V\rangle/k_1) + c_7$, $\beta_3 = c_6 \exp(f'_{c0}/E_0 - f'_c/E) \tanh(c_9\langle-\epsilon_V\rangle/k_1) + c_8$, $\alpha(\epsilon_I, \epsilon_{III}) = [k_5 / (1 + \min(\langle-\sigma_V^0\rangle, c_{21})/E_{N0})] \times [(\epsilon_I^0 - \epsilon_{III}^0)/k_1]^{c_{20}} + k_4$, and $\hat{\sigma}_N^0(\epsilon_V) = E_\tau(k_1c_{11} - c_{12}\langle\epsilon_V\rangle)$. In these equations, k_1, k_2, \dots, k_5 are the free parameters that need to be calibrated for any given concrete, while c_1, c_2, \dots, c_{20} are fixed constants for which the values are given in Table 1 of Caner and Bažant (2013b). Moreover, in tension the microplane normal modulus of elasticity degrades according to

$$E_N = E_{N0} \exp(-c_{13}\epsilon_N^{0+}) f(\xi) \tag{25}$$

where ϵ_N^{0+} is the maximum tensile fracturing strain ever reached on a microplane, a history variable; $f(\xi) = 1/(1 + 0.1\xi)$ where $\xi = \int \langle d\epsilon_V \rangle$ is a correction function to take into account the fatigue of concrete under tensile load cycles (Kirane and Bažant 2015). In compression, the microplane normal elastic modulus degradation occurs according to

$$E_N = E_{N0} [\exp(-c_{14}|\epsilon_N^{0-}|/[1 + c_{15}\epsilon_e^0]) + c_{16}\epsilon_e^0] \tag{26}$$

where ϵ_N^{0-} is the maximum compressive fracturing strain ever reached on a microplane, another history variable. In the foregoing equations; E_{N0} is the undamaged microplane normal elastic modulus, E_N is the (possibly damaged) microplane normal elastic modulus.

The elastic microplane stresses are calculated incrementally by

$$\begin{aligned}
\sigma_N^e &= \sigma_N^0 + E_N \Delta\epsilon_N \\
\tau_L^e &= \tau_L^0 + E_\tau \Delta\epsilon_L \\
\tau_M^e &= \tau_M^0 + E_\tau \Delta\epsilon_M \\
\tau^e &= \sqrt{\tau_L^{e2} + \tau_M^{e2}}
\end{aligned} \tag{27}$$

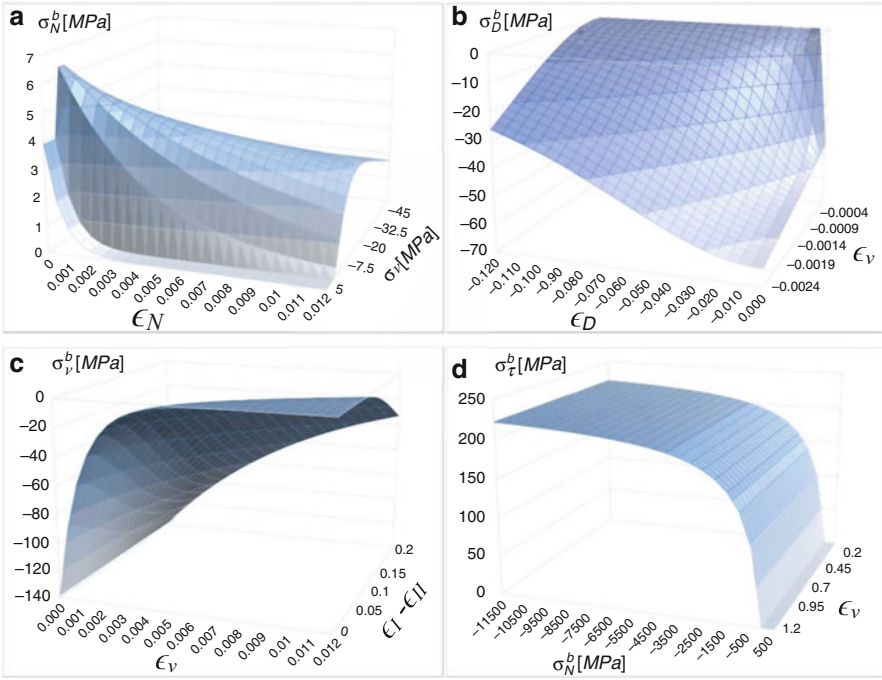


Fig. 2 The stress-strain boundaries of the model M7: (a) normal boundary, (b) deviatoric boundary, (c) volumetric boundary, and (d) friction boundary

where the microplane stresses from the previous load step σ_N^0 , τ_L^0 , and τ_M^0 are also history variables, leading to a total of five history variables per microplane. But the stress response on a microplane cannot exceed the corresponding stress-strain boundary:

$$\begin{aligned}
 \sigma_N &= \max \left(\min \left(\sigma_N^e, \mathcal{F}_N^+ \right), \mathcal{F}_N^- \right) \\
 \tau_L &= \frac{\tau_L^e}{\tau^e} \min \left(\tau^e, \mathcal{F}_\tau \right) \\
 \tau_M &= \frac{\tau_M^e}{\tau^e} \min \left(\tau^e, \mathcal{F}_\tau \right)
 \end{aligned}
 \tag{28}$$

Finally the second-order stress tensor is obtained by evaluating the integral in Eq. (6) using Gaussian quadrature. The strain-dependent yield limits of the model M7 given by the aforementioned equations are depicted in Fig. 2.

The experimental data first reported in Figs. 12 through 15 of Bažant et al. (1996a) features many different load paths in compression for the same concrete. To simulate these test data, once the model is calibrated, the model parameters cannot be changed from one load path prediction to the other, and thus these tests

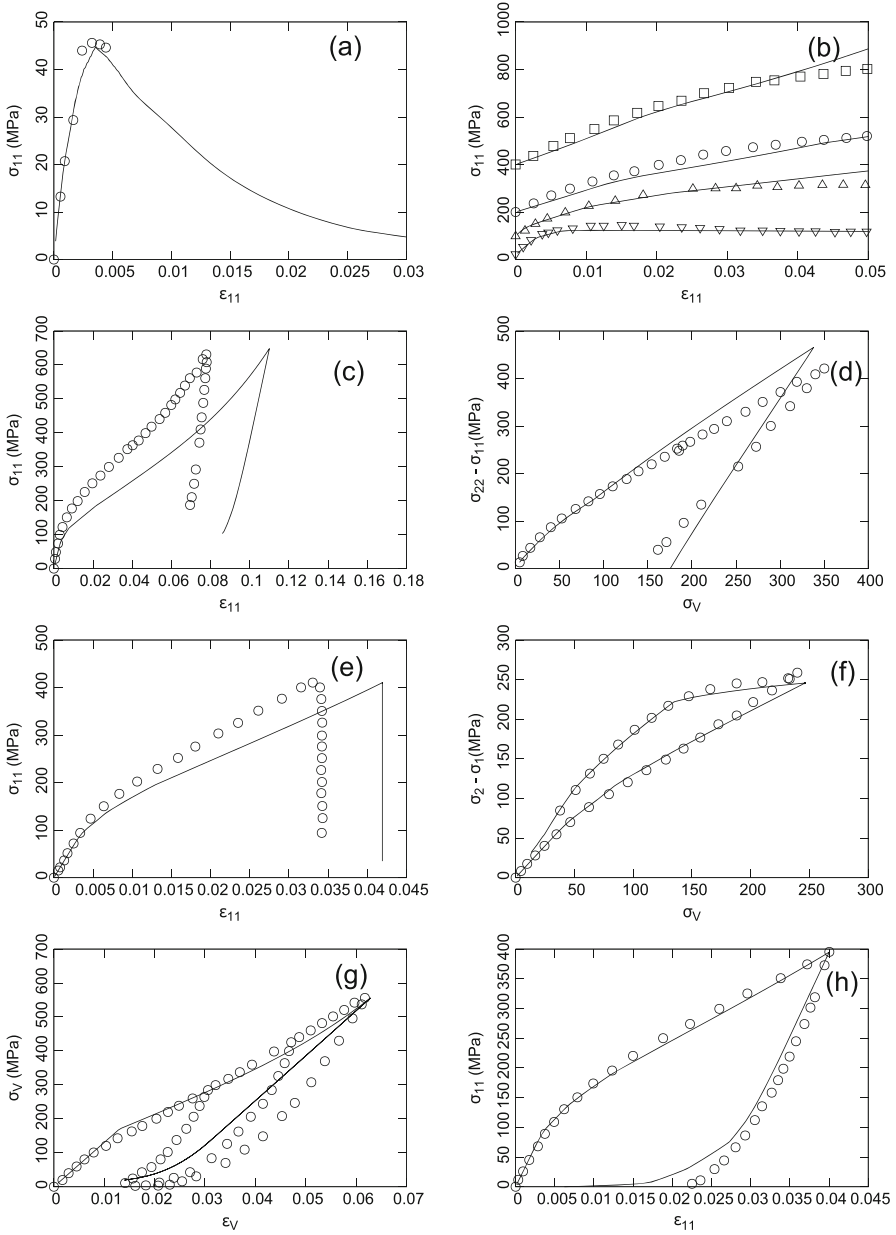


Fig. 3 The test data obtained at Waterways Experiment Station for different load paths using the same concrete (*the symbols*) and predictions by the model M7 (*the curves*): (a) uniaxial compression, (b) triaxial compression, (c, d) load path with lateral to axial strain ratio of -0.2 , (e, f) axial loading and lateral unloading, (g) hydrostatic compression, and (h) confined compression

are a major challenge for any concrete model. The model M7 successfully predicts response of this plain concrete at nearly all load paths as shown in Fig. 3 except for Fig. 3c. The test data for this load path for which the predicted response of the model M7 seems to be off could never be predicted correctly neither by the earlier microplane formulations nor by mesoscale models for plain concrete (see Cusatis et al. 2011) suggesting a likely problem with the test data. The performance of the model M7 in predicting the test data shown in these figures is about the same as that of the model M4. However, the model M7 calibration in this case also satisfies a most likely uniaxial tension behavior of this type of concrete, whereas the model M4 calibration did not take into account the tensile behavior of this concrete. Unfortunately a uniaxial tension test data for this concrete is not available, and thus the predicted response shown in this figure cannot be compared to the test data.

In Fig. 3 the parameter set of the model M7 employed for all simulations is given by $E = 25,000$ MPa, $\nu = 0.18$, $k_1 = 160.10^6$, $k_2 = 100$, $k_3 = 15$, $k_4 = 72$, $k_5 = 10.10^6$, $f'_{c0} = 15.08$ MPa, $E_0 = 20,000$ MPa, $c_1 = 0.089$, $c_2 = 0.4$, $c_3 = 0.4$, $c_4 = 50$, $c_5 = 3500$, $c_6 = 40$, $c_7 = 30$, $c_8 = 8$, $c_9 = 0.012$, $c_{10} = 0.4$, $c_{11} = 1.9$, $c_{12} = 0.18$, $c_{13} = 2500$, $c_{14} = 500$, $c_{15} = 7000$, $c_{16} = 100$, $c_{17} = 1$, $c_{18} = 1.6 \cdot 10^3$ MPa, $c_{19} = 1000$, $c_{20} = 1.8$, $c_{21} = 250$ MPa. Elements are deleted when the maximum microplane normal strain in tension exceeds 90% in all simulations.

A further extensive verification of the model M7 is given in Fig. 4. In these figures, the experimental data from three-point bending size effect tests originally published in Hoover et al. (2013) are simulated using the model M7. Once the model is calibrated using the load-displacement curve of the beam with 9.3 cm depth and a relative notch length of 30%, the model parameters remain fixed for the rest of the simulations. Thus, these size effect tests are a major challenge for any constitutive model for plain concrete. As shown in these figures, the performance in tension of the model M7 is outstanding. In addition, the model M7 can predict compression-tension load cycles unlike the earlier formulations of microplane models as shown in Fig.4 of Caner and Bažant (2013b) and in Kirane and Bažant (2015).

Furthermore, in Fig. 5 the compression test data published in Wendner et al. (2015) are predicted using the model M7. The test data given in Fig. 5a is for a cylindrical specimen aged for 31 days, and the data given in Fig. 5b is for a cylindrical specimen aged for 400 days. Other than this difference, the two concretes come from the same batch as the specimens of Fig. 4. The test data of Fig. 5c, d are for cubic specimens aged for 470 days. The compressive response parameters of the model M7 are calibrated in Fig. 5c, and for each case, the reported value of elastic modulus is employed. In addition, in all compression simulations, the contact conditions between test platens and the specimens are simulated using a band of cohesive elements. These contact conditions turn out to be critical to obtain the result shown in Fig. 5d. The tensile response of the model remained the same as in Fig. 4. As can be inferred from the fits in these figures, the model M7 can reproduce the compressive response of concrete to a large extent simultaneously with the tensile response using the same set of parameters.

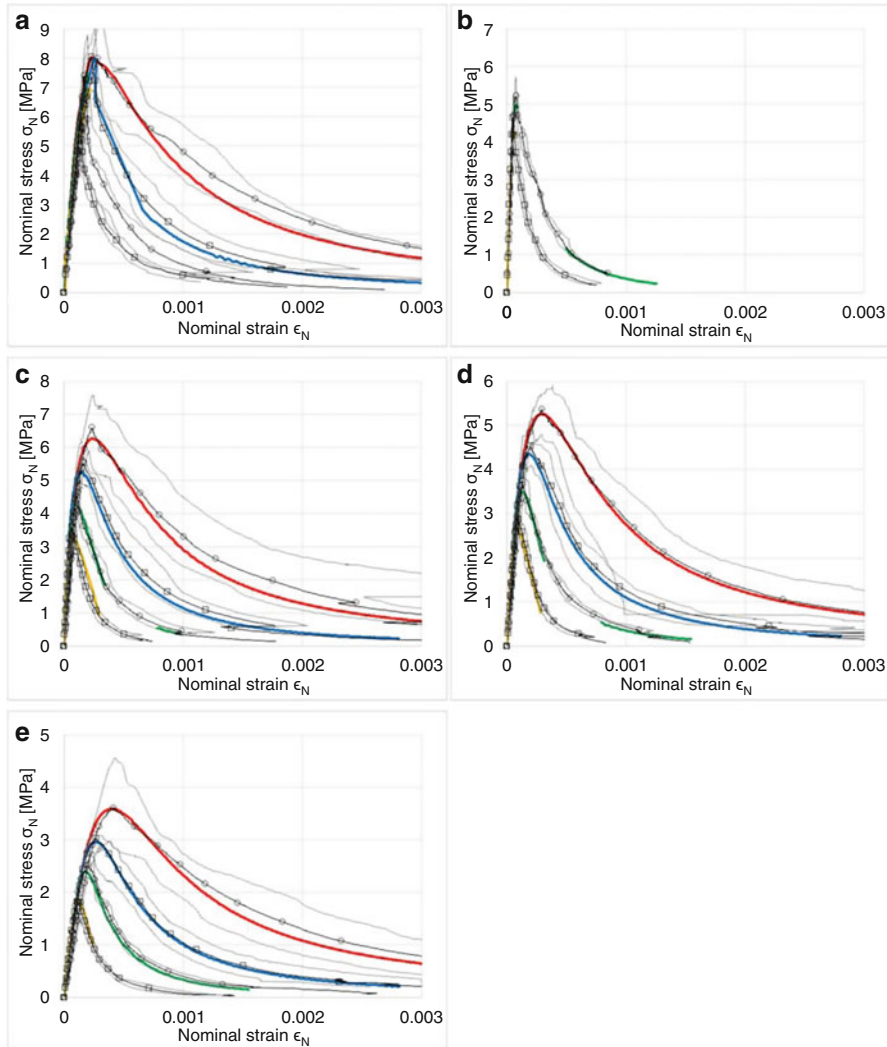


Fig. 4 Curves with symbols show the size effect test data from three-point bending tests of plain concrete specimens with various notch lengths relative to beam depth: **(a)** 0% notch depth (unnotched), **(b)** 2.5% notch depth, **(c)** 7.5% notch depth, **(d)** 15% notch depth, and **(e)** 30% notch depth. Light colored thin lines show the range of test data. The predictions by the model M7 are shown by thick curves in color

The specimens with predicted cracks are shown in Fig. 6. The cylindrical specimens exhibit a diagonal band of axial splitting cracks at failure. The 4 cm cubic specimen shown in Fig. 6c develops extensive distributed splitting cracks at failure. In contrast, the major mode of failure in the 15 cm cubic specimen

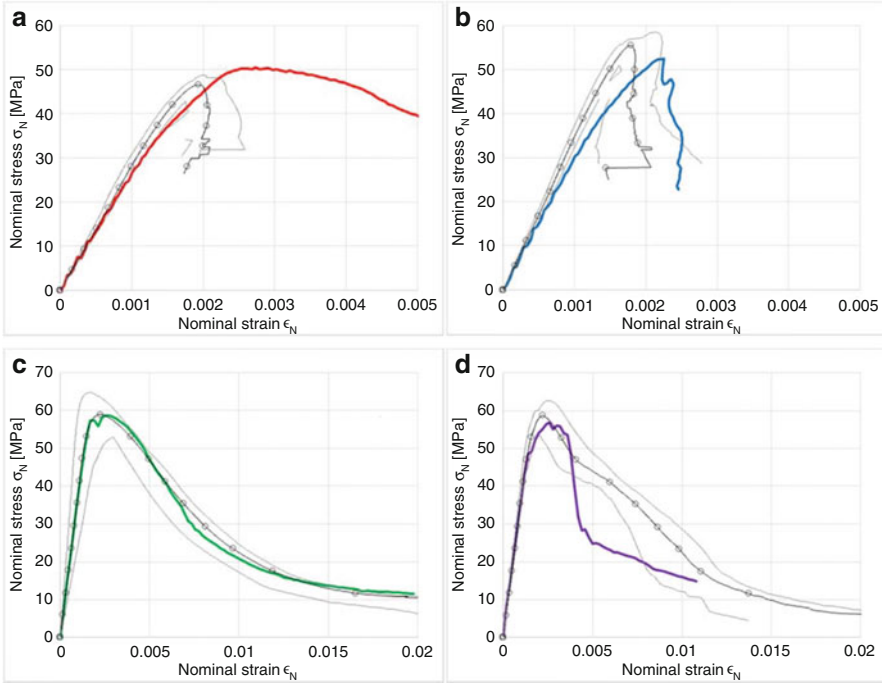


Fig. 5 Compression test data obtained from the specimens from an overcured version of the same concrete as the three-point bending plain concrete specimens of Fig. 4: (a) Cylindrical specimen of concrete aged 31 days with dimensions 7.5 by 15 cm, (b) cylindrical specimen of concrete aged 400 days with the same dimensions, (c) cubic specimen with a dimension of 4 cm, and finally (d) cubic specimen with a dimension of 15 cm. The *curve with symbols* is the test data; *light colored thin lines* show the range of test data. The predictions by the model M7 are shown by *thick curves in color*

shown in Fig. 6d is a large splitting crack in the middle. The hexahedral mesh used in all these calculations of compression failure has an average size of 5 mm.

In Figs. 4 through 6 the parameter set of the model M7 employed for all simulations is given by $\nu = 0.18, k_1 = 125 \cdot 10^6, k_2 = 15, k_3 = 15, k_4 = 20, k_5 = 16 \cdot 10^6, f'_{c0} = 15.08$ MPa, $E_0 = 20,000$ MPa, $c_1 = 0,089, c_2 = -1.318, c_3 = 8.4, c_4 = 70, c_5 = 3500, c_6 = 20, c_7 = 80, c_8 = 15, c_9 = 0.012, c_{10} = 1, c_{11} = 0.5, c_{12} = 2.36, c_{13} = 4500, c_{14} = 300, c_{15} = 4000, c_{16} = 60, c_{17} = 1.4, c_{18} = 17.5$ MPa, $c_{19} = 14,000, c_{20} = 1.8, c_{21} = 250$ MPa. Elements are deleted when the maximum microplane normal strain in tension exceeds 90%. The Young modulus employed to calculate the curves of Fig. 4 is given by $E = 41,290$ MPa. In Figs. 5a and 6a $E = 27,735$ MPa, in Figs. 5b and 6b $E = 31,382$ MPa, in Figs. 5c and 6c $E = 40,509$ MPa and in Figs. 5d and 6d $E = 38,063$ MPa are used.

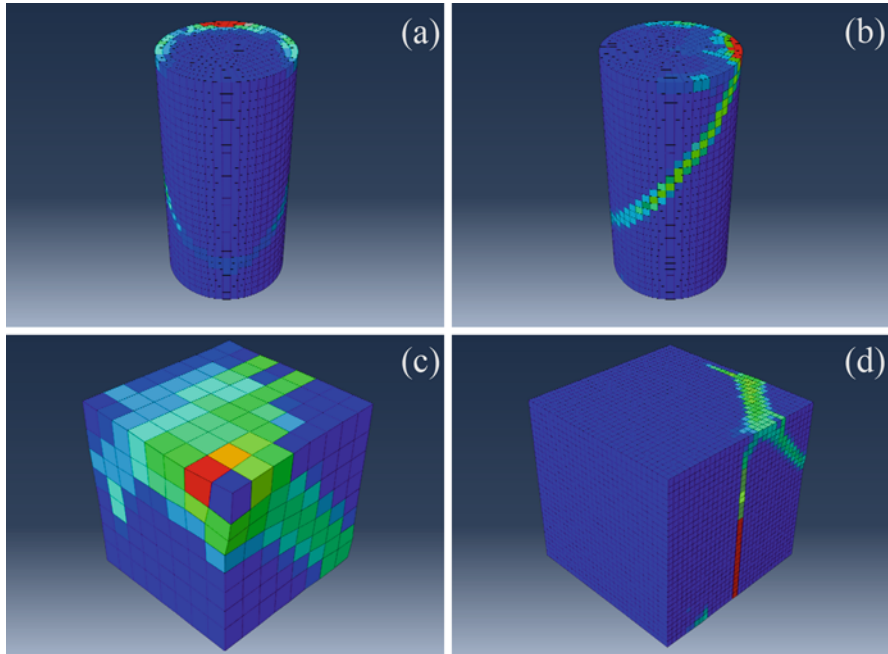


Fig. 6 (a) Cylindrical specimen of concrete aged 31 days with dimensions 7.5 by 15 cm, (b) cylindrical specimen of concrete aged 400 days with the same dimensions, (c) cubic specimen with a dimension of 4 cm, and finally (d) cubic specimen with a dimension of 15 cm

Nonlocal Microplane Models

In the analysis of structures by the finite element method, the microplane model is meant to be used to calculate the second-order stress tensor that corresponds to a given second-order strain tensor at the Gauss points of a finite element mesh made of continuum elements. The fracture predicted by the microplane model at the Gauss points is distributed across the width of an element creating a crack band. This approach to modeling fracture has a number of advantages including automatic tracking of crack propagation direction and automatic detection of crack nucleation anywhere in the mesh. It also allows the use of the well-developed continuum mechanics theories in the simulation of fracture and damage. However, one disadvantage to crack band modeling is that the calculated response depends on the width of the crack band. The smaller the width, the smaller is the energy dissipated by fracture; thus in the limit of vanishing element width, fracture energy also vanishes. Obviously this is unrealistic. There seems to be three choices to get the correct energy dissipation as the cracks propagate in the material: (1) the element width may be considered as a material parameter in which case, the finite element mesh should consist of elements of one single size and one single aspect ratio

close to that of a cube for a general 3-D fracture propagation and the microplane model should be calibrated for that element size, (2) a localization limiter may be introduced by which the average of the variables of the model that govern the fracturing of the material or equivalently of the fracturing part of the total strains over a finite neighborhood of every Gauss point may be calculated, and finally (3) using principle of virtual work, the strongly nonlocal partial differential stress equations of motion with a characteristic length which also yields the corresponding boundary conditions may be developed.

The nonlocal integral generalization of the microplane split-type formulation given in Bažant and Prat (1988a) is presented in Bažant and Ožbolt (1990) as the first ever nonlocal microplane model. In this formulation, the localization of the fracture process zone into zero width as the mesh is refined is prevented by introducing the volume averages of only damage variables, equivalent to averaging the fracturing part of the total strains. Thus, the elastic part of the response remains local. The calculation of the volume average when part of the volume protrudes off the boundary, the integration weights are rescaled so as to produce a true average. This kind of treatment of the boundary zone is one of the reasons for the stiffness matrix to become nonsymmetric. The weight function employed is a bell-shaped one, similar to the Gaussian probability distribution function, with a cutoff distance beyond which the value of the function becomes zero. The fact that many other weight functions can produce satisfactory fits of experimental data means a unique weight function does not exist. This formulation leads to a mesh size-independent model (excluding the discretization error that results when coarse meshes are used). In Bažant and Ožbolt (1992), the nonlocal formulation presented in Bažant and Ožbolt (1990) is employed to simulate the compression fracture of concrete which is more complex than the tensile fracture of concrete.

The nonlocal formulation presented in Bažant and Di Luzio (2004) extends the microplane model M4 given in Bažant et al. (2000) to nonlocality. The over-nonlocal formulation in which the strains in the stress-strain boundaries of the model M4 are replaced by the total microplane strains averaged over a finite neighborhood of the Gauss point, multiplied by an empirical coefficient m and added to total microplane strain multiplied by $(1-m)$, is presented in di Luzio (2007) and Di Luzio and Bažant (2005). This formulation has the advantage that it does not require a variable that governs the fracture nor the separation of the total microplane strains into fracturing and elastic parts. However, it also suffers from the typical problems of the classical integral type nonlocal formulations.

A Mesh Size-Independent Microplane Model: The Model M7Auto

Although the nonlocal integral type crack band models achieve the mesh size independence of the calculated response, they have a number of disadvantages and problems. For example, the nonlocal crack band models cannot capture the closely spaced distributed cracking due to either the reinforcement in a deforming reinforced concrete beam or the drying shrinkage of concrete surface, whereas the

mesh size-dependent crack band models can indeed capture those cracks accurately provided that they are calibrated for the element size employed in the mesh. Another important problem in the nonlocal crack band modeling is the treatment of the boundary. When the averaging volume protrudes the boundary of the structure, the averaging volume has to be reduced, which means that at the boundary, the averaging becomes unobjective. The least problematic treatment of the boundary in the sense of nonlocal crack band modeling is probably keeping a sufficiently thick boundary layer where the model is kept mesh size *dependent* with a calibration for that mesh size employed in the boundary layer (Bažant et al. 2010). Yet another problem in such nonlocal crack band models is that cracks initiate ahead of the crack tip which is not physically realistic. Furthermore, in dynamic applications, the separation of the particles smaller than the averaging size cannot be captured. Moreover, there is no practical way to detect the shielding effect of a partially formed cohesive crack between two Gauss points within the same averaging volume. Yet another difficulty is that a varying size of the averaging volume is needed for correctly capturing crack initiation and crack growth. A computational difficulty in Abaqus (Simulia Corporation 2014) is that for large meshes, it is not possible to correctly calculate the volume averages of the relevant fracturing variable of a crack band model because the user subroutine is evaluated in separate blocks of Gauss points.

However, if the continuum model can calibrate itself for different element size, aspect ratio, and type, not only the model response becomes independent of mesh size, but also all the aforementioned difficulties associated with classical nonlocal integral type approach could be avoided. Thus, in what follows, an autocalibrating version of the model M7 called the model M7Auto is described, and its performance is illustrated using several examples.

In the model M7, the softening behavior in tension and in compression is governed by the parameters c_4 , c_{11} , c_{12} , and c_{13} . Thus, with varying element size, these four fixed parameters must vary so as to dissipate the same fracture energy. The resulting element size-dependent stress-strain boundaries are obtained by substituting

$$\begin{aligned} \mathcal{F}_N^+(\epsilon_N, \sigma_V) &= Ek_1\beta_1 \exp\left(-\langle\epsilon_N - \beta_1 c_2 k_1\rangle / \left(\left(c_4 + \frac{s_9 - c_4}{\left(1 + \left(\frac{l_{ch}}{s_{10}}\right)^{s_{11}}\right)^{s_{12}}}\right)\epsilon_e^0 + k_1 c_3\right)\right) \\ \hat{\sigma}_N^0(\epsilon_V) &= E_\tau \left\langle k_1 \left(c_{11} + \frac{s_1 - c_{11}}{\left(1 + \left(\frac{l_{ch}}{s_2}\right)^{s_3}\right)^{s_4}}\right) - \left(c_{12} + \frac{s_5 - c_{12}}{\left(1 + \left(\frac{l_{ch}}{s_6}\right)^{s_7}\right)^{s_8}}\right) \langle\epsilon_V\rangle \right\rangle \quad (29) \\ E_N &= E_{N0} \exp\left(- (112l_{ch} + 1780)\epsilon_N^{0+}\right) f(\xi) \end{aligned}$$

in Eqs. (24) and (25). In Figs. 7 through 10 the parameter set of the model M7 employed for all simulations is given by $E = 30,500$ MPa, $\nu = 0.18$, $k_1 =$

$189 \cdot 10^6$, $k_2 = 110$, $k_3 = 24$, $k_4 = 6$, $k_5 = 100 \cdot 10^6$, $f'_{c0} = 15.08$ MPa, $E = 20,000$ MPa, $c_1 = 0,089$, $c_2 = -1.9$, $c_3 = 3.7538$, $c_4 = 25.85335$, $c_5 = 3500$, $c_6 = 20$, $c_7 = 1$, $c_8 = 8$, $c_9 = 0.012$, $c_{10} = 0.33$, $c_{11} = 184,434.5$, $c_{12} = 264,009.5$, $c_{13} = 112 l_{ch} + 1780$, $c_{14} = 300$, $c_{15} = 4000$, $c_{16} = 60$, $c_{17} = 1.4$, $c_{18} = 1.6 \cdot 10^{-3}$ MPa, $c_{19} = 1000$, $c_{20} = 1.8$, $c_{21} = 250$ MPa. The newly introduced “s” parameters are given by $s_1 = 0.981856$, $s_2 = 21.67212$, $s_3 = 4.291001$, $s_4 = 1.51 \cdot 10^{-6}$, $s_5 = 0.020228$, $s_6 = 16.85614$, $s_7 = 7.696828$, $s_8 = 1.8 \cdot 10^{-6}$, $s_9 = 124.9338$, $s_{10} = 7,625,177$, $s_{11} = 1.15299$ and $s_{12} = 3,048,460$. Elements are deleted when the maximum microplane normal strain in tension exceeds 90% in all simulations. The parameter l_{ch} is the characteristic length of the element that corresponds to the variable “charLength” in Abaqus user subroutine. The reference fracture energy can be obtained by calibrating the model using a mesh with a fixed element size, type, and aspect ratio close to 1. During this initial calibration phase, the free parameters of the model k_1 through k_5 also need to be varied. Next, a new mesh is generated with the desired element size keeping the element type the same and the aspect ratio as close to 1 as possible. Now the free parameters are kept fixed, and the four fixed parameters are varied until the same fracture energy is dissipated both in tension and in uniaxial and triaxial compression. This procedure is repeated for different element sizes, and the resulting trends in those four fixed parameters are approximated using appropriate functions. These variations of the fixed parameters as functions of element size are programmed into the model M7 in the form of a user subroutine for Abaqus. This allows the model to recalibrate itself for different element sizes within a range of 5 to 20 mm. The upper limit of this range can further be increased using the approach given in Červenka et al. (2005) but that requires a new user element to be defined in Abaqus.

In concrete the characteristic size is considered to be roughly the half of the size of the fracture process zone. In crack band models, the width, w , of the crack band corresponds to the size of the fracture process zone and thus is a measure of the characteristic size of the material. In the finite element calculations, it seems reasonable to use elements of size about three to four times the maximum aggregate size of the concrete of interest, and it seems unlikely that smaller elements can be used to simulate the softening behavior. However, as shown in this study, much smaller crack band widths can be employed in the finite element calculations to obtain correct material response provided that the model is properly calibrated for these widths. Thus, for a different element size w , the predicted energy per unit volume dissipated by fracture, γ_f , should vary in order to satisfy

$$G_f = w\gamma_f \quad (30)$$

where G_f is constant in quasi-static loading conditions (see Fig. 1 in Bažant and Caner 2005). Although it is found out in this study that for much smaller values of w than three to four times the maximum aggregate size the model can be successfully recalibrated in order to satisfy Eq. (30), when the width of the element (or the crack

band) is much larger, satisfying this equation becomes very difficult and a different strategy as in Červenka et al. (2005) may be needed.

In Abaqus documentation (Simulia Corporation 2014), the precise calculation of the characteristic length of an element in three-dimensional space is not given. In the finite element calculations, however, it is found out that for both bilinear hexahedral and linear tetrahedral elements, the characteristic length made available to the user subroutine as the variable “charLength” coincides with the cube root of the element volume:

$$\text{charLength} = \sqrt[3]{V} \quad (31)$$

For a hexahedral element with an aspect ratio of 1, the characteristic length is the same as its side length. This is convenient for the model calibration in problems in which the cracks propagate along the mesh lines in a hexahedral mesh using Abaqus. For quadratic tetrahedral elements, the characteristic length is given by

$$\text{charLength} = \sqrt[3]{V/4} \quad (32)$$

The tetrahedral elements not only allow meshing of more complex geometries compared to hexahedral elements, but also in crack band models, they pose less resistance to a possible change in the direction of cracks, and thus potentially they are more useful than hexahedral elements. However in this study both hexahedral structured meshes and tetrahedral random meshes have been employed in the calibration of the model. For the case of a priori knowledge of crack paths, the hexahedral elements are preferred because the characteristic size is more closely related to element dimensions in these elements than in the tetrahedral elements.

In the following finite element analyses, the aspect ratio of the elements is kept as close to 1 as possible. For an arbitrary geometry, even for a geometry as simple as a cylindrical one, it is not trivial to obtain a mesh made of elements with an aspect ratio very close to unity. However, the aspect ratio of the elements must be as close to unity as possible to be able capture the correct energy dissipation as well as the correct crack paths.

The Nooru-Mohamed-mixed mode fracture test data set presented in Nooru-Mohamed (1992) for different load paths is a well-known challenge for constitutive models for plain concrete. The finite element model required to fit these test data involves a correct representation of the test setup including the rigidities of the three frames used in the tests in addition to the other restrictions related to finite element mesh used in the sense of crack band model to discretize the concrete specimens. First, the model M7Auto is calibrated by optimally fitting the experimental load-displacement data with a mesh oriented along the experimentally observed crack paths (Fig. 7d). Next, the performance of the model M7Auto is tested by employing finite element random meshes with different element sizes and types. The load-displacement predictions are compared to the test data in Fig. 7f. In this figure, the dark-blue curve is associated with the mesh in Fig. 7a, the thick gray curve

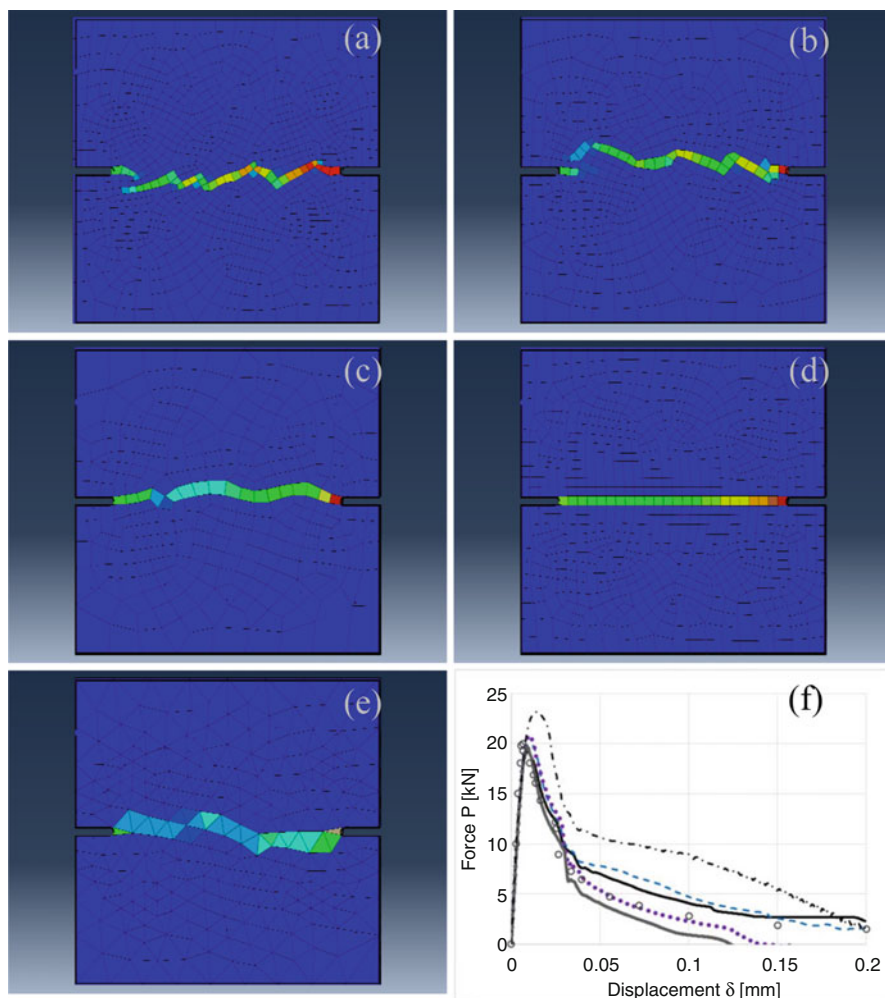


Fig. 7 Crack paths and load displacement curves predicted by the model M7Auto in uniaxial tension of the Nooru-Mohamed specimens using different mesh sizes: (a) 5 mm hexahedral random mesh, (b) 6.25 mm hexahedral random mesh, (c) 10 mm hexahedral random mesh, (d) 6.25 mm hexahedral mesh oriented along the experimentally observed crack paths, (e) 5 mm random tetrahedral mesh, (f) the load-displacement experimental data shown as *symbols* and the predictions by the model M7Auto

with the mesh of Fig. 7d, the dotted purple curve with the mesh of Fig. 7c, the dashed blue curve with the mesh in Fig. 7b, and finally the blue dashed dotted curve with the mesh in Fig. 7e. The crack paths predicted in this simple load path are in general reasonable, and the load-displacement predictions agree very well with the experimental data with the exception of the case of Fig. 7e in which relatively large tetrahedral elements are employed.

In Fig. 8a–e, the experimentally observed crack patterns that correspond to the load path 4a and the crack paths predicted by the already calibrated model M7Auto are shown. The average mesh sizes used in the calculations are approximately 5, 6.25, and 10 mm as shown in the figures. The meshes are random in Fig. 8a, b, c, e, whereas in Fig. 8d the mesh is oriented along the experimentally observed crack paths. Clearly the crack paths in Fig. 8d are captured almost exactly by the model as the cracks do not have to propagate skew to the mesh in a zigzag pattern. In Fig. 8a, b, c, e the crack paths are still captured reasonably well, but especially more accurately when the element size is small enough. In Fig. 8e a random mesh of linear tetrahedral elements are used, and comparing this figure to Fig. 8a in which bilinear hexahedral elements are used, it is observed that tetrahedral elements prevent the zigzag pattern of crack propagation skew to the mesh. The predicted load-displacement response corresponding to Fig. 8a–e and test data are given in Fig. 8f. In this figure, the dark-blue curve is associated with the mesh in Fig. 8a, the thick gray curve with the mesh of Fig. 8d, the dotted purple curve with the mesh of Fig. 8c, the dashed blue curve with the mesh in Fig. 8b, and finally the blue dashed dotted curve with the mesh in Fig. 8e. The peak load is generally overestimated, but the response predicted by the model M7Auto in general agrees well with the test data with the exception of the case of Fig. 8e.

Figure 9a–e show the crack patterns that result from the load path 4b and the crack paths predicted by the model M7Auto. Similar to Fig. 8, the average mesh sizes are approximately 5, 6.25, and 10 mm in these figures. The meshes are random in Fig. 9a, b, c, e, whereas in Fig. 9d the mesh is oriented along the experimentally observed crack paths. Clearly the crack paths in Fig. 9d are simulated very well by the model as the cracks do not have to propagate skew to the mesh. In Fig. 9a, b, c, e the crack paths are still simulated reasonably well, but the accuracy is higher when the element size is small enough. In Fig. 9e a random mesh of linear tetrahedral elements are used. Comparing this figure to Fig. 9a in which bilinear hexahedral elements are used, it is observed that tetrahedral elements facilitate the crack propagation skew to the mesh. The predicted load-displacement response corresponding to Fig. 9a–e and test data are shown in Fig. 9f. In this figure, the dark-blue curve corresponds to the mesh in Fig. 9a, the thick gray curve to the mesh of Fig. 9d, the dotted purple curve to the mesh of Fig. 9c, the dashed blue curve to the mesh in Fig. 9b, and finally the blue dashed dotted curve to the mesh in Fig. 9e. The peak load is again mostly overestimated, but the response predicted by the model M7Auto in general is in a reasonable agreement with the test data although the case of Fig. 9e shows a significantly large deviation from the test data.

Similarly, Fig. 10a–e show the experimentally observed crack patterns that correspond to the load path 4c and the crack paths predicted by the already calibrated model M7Auto. Similar to Fig. 8, the average mesh sizes shown are approximately 5, 6.25, and 10 mm in these figures. The meshes are random in Fig. 10a, b, c, e, whereas in Fig. 10d the mesh is oriented along the experimentally observed crack paths. Clearly the crack paths in Fig. 10d are captured almost exactly by the model as the cracks do not have to propagate skew to the mesh in a zigzag pattern. In Fig. 10a, b, c, e the crack paths are still captured

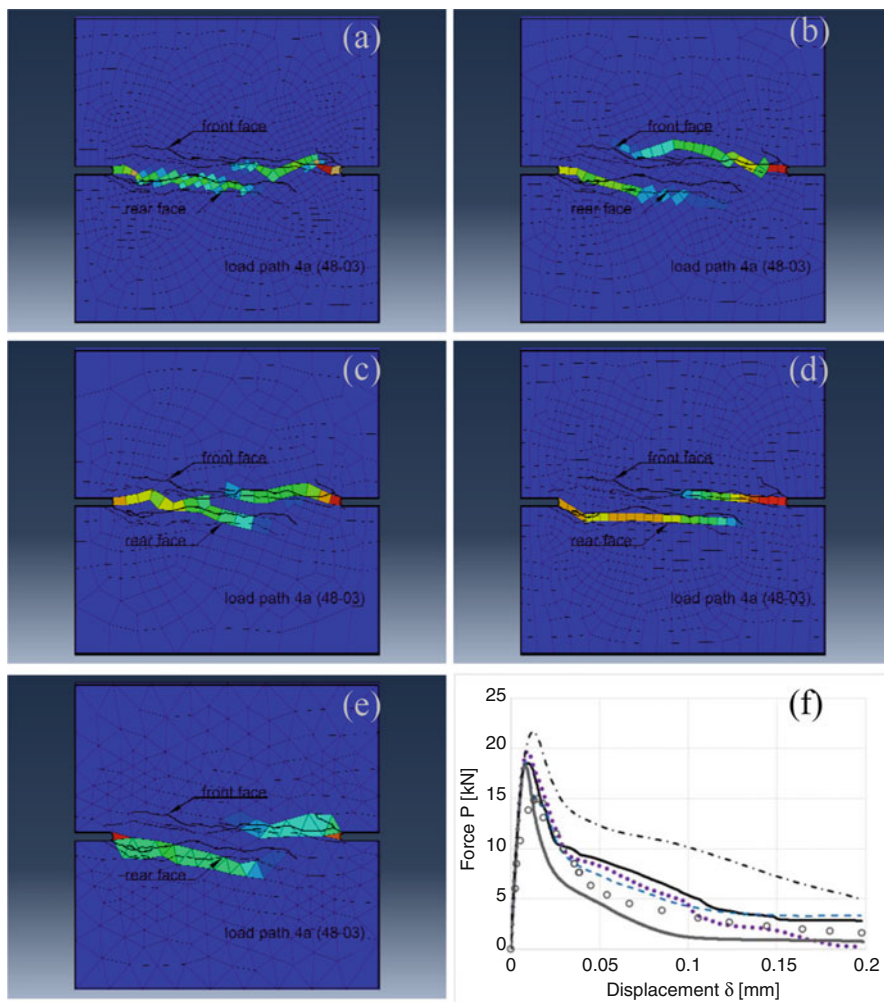


Fig. 8 Crack paths and load displacement curves predicted by the model M7Auto as the response of the Nooru-Mohamed specimens to the load path 4a using different mesh sizes: (a) 5 mm hexahedral random mesh, (b) 6.25 mm hexahedral random mesh, (c) 10 mm hexahedral random mesh, (d) 6.25 mm hexahedral mesh oriented along the experimentally observed crack paths, (e) 5 mm random tetrahedral mesh, (f) the load-displacement experimental data shown as *symbols* and the predictions by the model M7Auto

reasonably well, but especially more accurately when the element size is small enough. In Fig. 10e a random mesh of linear tetrahedral elements are used, and comparing this figure to Fig. 10a in which bilinear hexahedral elements are used, it is observed that tetrahedral elements prevent the zigzag pattern of crack propagation skew to the mesh. The predicted load-displacement response corresponding to Fig. 10a–e and test data are given in Fig. 10f. In this figure,

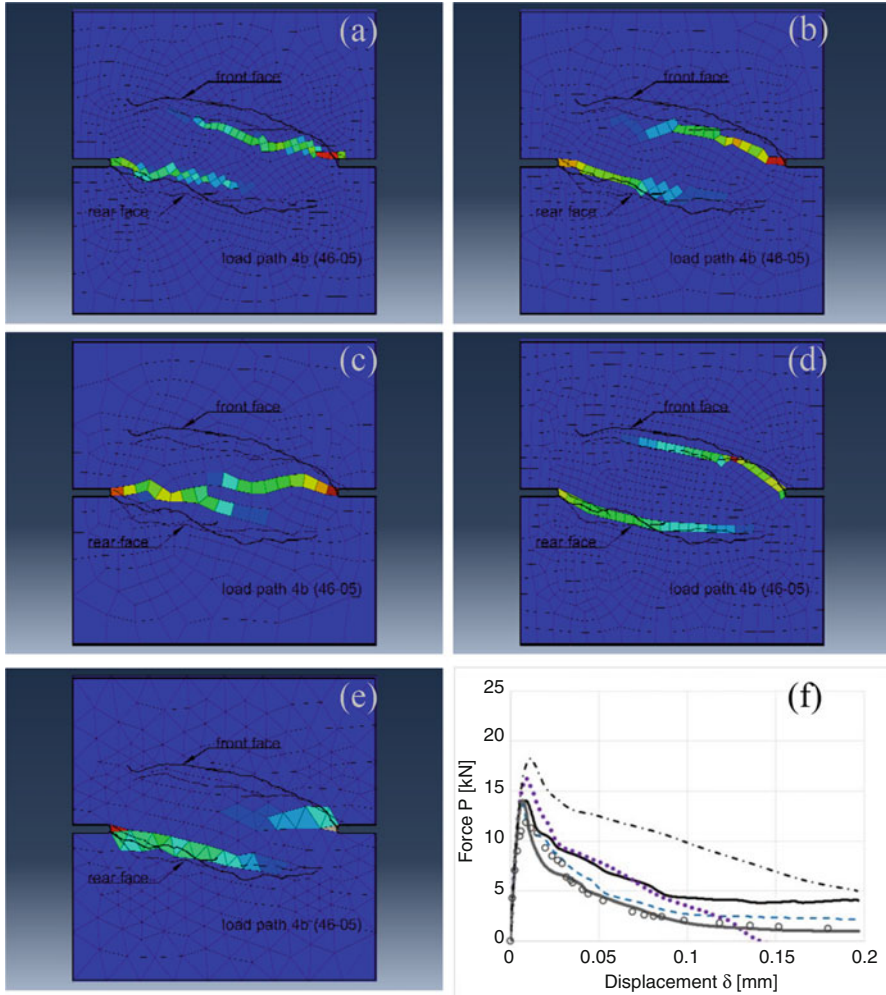


Fig. 9 Crack paths and load displacement curves predicted by the model M7Auto as the response of the Nooru-Mohamed specimens to the load path 4b using different mesh sizes: (a) 5 mm hexahedral random mesh, (b) 6.25 mm hexahedral random mesh, (c) 10 mm hexahedral random mesh, (d) 6.25 mm hexahedral mesh oriented along the experimentally observed crack paths, (e) 5 mm random tetrahedral mesh, (f) the load-displacement experimental data shown as *symbols* and the predictions by the model M7Auto

the dark-blue curve is associated with the mesh in Fig. 10a, the thick gray curve with the mesh of Fig. 10d, the dotted purple curve with the mesh of Fig. 10c, the dashed blue curve with the mesh in Fig. 10b, and finally the blue dashed dotted curve with the mesh in Fig. 10e. The peak load is generally overestimated, but the response predicted by the model M7Auto in general agrees well with the test data with the exception of the case of Fig. 10e.

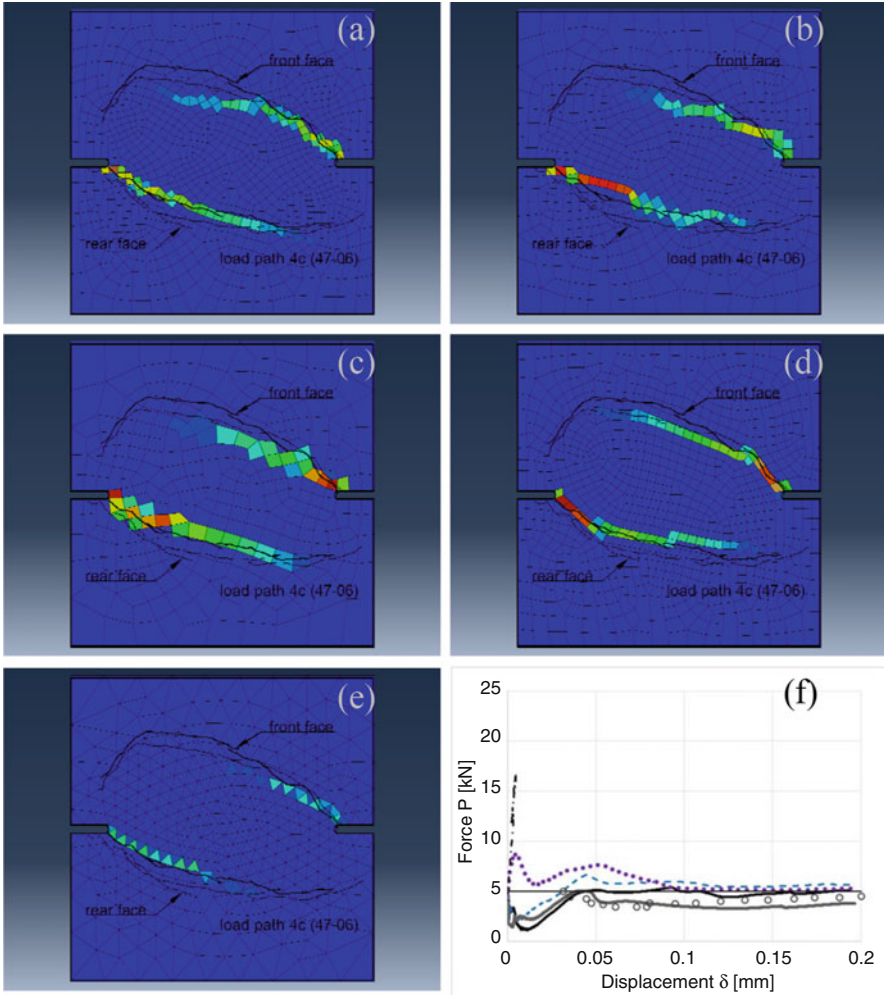


Fig. 10 Crack paths and load displacement curves predicted by the model M7Auto as the response of the Nooru-Mohamed specimens to the load path 4c using different mesh sizes: (a) 5 mm hexahedral random mesh, (b) 6.25 mm hexahedral random mesh, (c) 10 mm hexahedral random mesh, (d) 6.25 mm hexahedral mesh oriented along the experimentally observed crack paths, (e) 5 mm random tetrahedral mesh, (f) the load-displacement experimental data shown as symbols and the predictions by the model M7Auto

Although the predictions corresponding to the cases given in Figs. 7e, 8e, 9e, and 10e seem to be too inaccurate, in fact the coarse random mesh can cause this kind of deviations in response due to discretization error; see Figs. 5 and 6 in Červenka et al. (2005). These calculation results reinforce the idea that if the crack paths are not known in advance, a two-step calculation procedure should be followed: in the trial step, the calculations should be carried out using a random mesh, and in the

final step, the calculations should be done using a mesh oriented along the crack paths observed in the trial step.

Conclusions

In this study three types of common microplane formulations that can be applied to a wide range of material constitutive behavior are discussed. The focus is on the microplane models for plain concrete, a challenging engineering material to model mathematically. Only the major milestones in the development of microplane models for concrete are explicitly treated. Thus, the model M4, the model M7, and a novel nonlocal extension of the model M7 called the model M7Auto are discussed in this study. The predictive capabilities of the model M7 are demonstrated by simulating the size effect test data obtained by three-point bending tests of geometrically similar beams as well as the compression test data obtained using specimens from an overcured version of the same concrete at different proportional and nonproportional load paths. Thus, it is shown that the model M7 fits all these experimental data very accurately.

Furthermore, it is shown that the model M7Auto overcomes the difficulties associated with classical approaches to nonlocal modeling simply by recalibrating itself according to the element size, shape, and type. The requirements for a successful use of the model M7Auto are that (1) the size of the elements in the mesh should not exceed the range of autocalibration of the model, (2) the elements must have an aspect ratio close to 1 for the simulation of general three-dimensional crack propagation, (3) the element type should be the same in the mesh. The predictive capabilities of the model M7Auto are demonstrated by simulating some well-known load paths in the benchmark test data of Nooru-Mohamed. It is shown that the model M7Auto can handle a significant range of element sizes from 5 to 20 mm and still predict the concrete response to complex load paths quite accurately.

References

- Z. Bažant, M. Adley, I. Carol, M. Jirasek, S. Akers, B. Rohani, J. Cargile, F. Caner, Large-strain generalization of microplane model for concrete and application. *J. Eng. Mech. ASCE* **126**(9), 971–980 (2000)
- Z. Bažant, F. Caner, Microplane model M5 with kinematic and static constraints for concrete fracture and anelasticity. II: computation. *J. Eng. Mech. ASCE* **131**(1), 41–47 (2005)
- Z. Bažant, F. Caner, M. Adley, S. Akers, Fracturing rate effect and creep in microplane model for dynamics. *J. Eng. Mech. ASCE* **126**(9), 962–970 (2000)
- Z. Bažant, F. Caner, I. Carol, M. Adley, S. Akers, Microplane model M4 for concrete. I: formulation with work-conjugate deviatoric stress. *J. Eng. Mech. ASCE* **126**(9), 944–953 (2000)
- Z. Bažant, G. Di Luzio, Nonlocal microplane model with strain-softening yield limits. *Int. J. Solids Struct.* **41**(24–25), 7209–7240 (2004)
- Z. Bažant, P. Gambarova, Crack shear in concrete – crack band microplane model. *J. Eng. Mech. ASCE* **110**(9), 2015–2035 (1984)

- Z. Bažant, J.-L. Le, C. Hoover, Nonlocal boundary layer (NBL) model: overcoming boundary condition problems in strength statistics and fracture analysis of quasibrittle materials, in *Proceedings of FramCoS-7*, 2010, pp. 135–143
- Z. Bažant, B. Oh, Microplane model for progressive fracture of concrete and rock. *J. Eng. Mech. ASCE* **111**(4), 559–582 (1985)
- Z. Bažant, J. Ožbolt, Nonlocal microplane model for fracture, damage, and size effect in structures. *J. Eng. Mech. ASCE* **116**(11), 2485–2505 (1990)
- Z. Bažant, J. Ožbolt, Compression failure of quasibrittle material – nonlocal microplane model. *J. Eng. Mech. ASCE* **118**(3), 540–556 (1992)
- Z. Bažant, G. Pijaudiercabot, Nonlocal continuum damage, localization instability and convergence. *J. Appl. Mech. Trans. ASME* **55**(2), 287–293 (1988)
- Z. Bažant, P. Prat, Microplane model for brittle-plastic material. 1. Theory. *J. Eng. Mech. ASCE* **114**(10), 1672–1687 (1988a)
- Z. Bažant, P. Prat, Microplane model for brittle-plastic material. 2. Verification. *J. Eng. Mech. ASCE* **114**(10), 1689–1702 (1988b)
- Z. Bažant, Y. Xiang, M. Adley, P. Prat, S. Akers, Microplane model for concrete. 2. Data delocalization and verification. *J. Eng. Mech. ASCE* **122**(3), 255–262 (1996a)
- Z. Bažant, Y. Xiang, P. Prat, Microplane model for concrete. 1. Stress-strain boundaries and finite strain. *J. Eng. Mech. ASCE* **122**(3), 245–254 (1996b)
- Z. Bažant, G. Zi, Microplane constitutive model for porous isotropic rocks. *Int. J. Numer. Anal. methods Geomech.* **27**(1), 25–47 (2003)
- M. Brocca, Z. Bažant, Microplane constitutive model and metal plasticity. *Appl. Mech. Rev. ASME* **53**(10), 265–281 (2000)
- M. Brocca, Z. Bažant, I. Daniel, Microplane model for stiff foams and finite element analysis of sandwich failure by core indentation. *Int. J. Solids Struct.* **38**(44-45), 8111–8132 (2001)
- M. Brocca, L. Brinson, Z. Bažant, Three-dimensional constitutive model for shape memory alloys based on microplane model. *J. Mech. Phys. Solids* **50**(5), 1051–1077 (2002)
- F. Caner, Z. Bažant, Microplane model M4 for concrete. II: algorithm and calibration. *J. Eng. Mech. ASCE* **126**(9), 954–961 (2000)
- F. Caner, Z. Bažant, J. Cervenka, Vertex effect in strain-softening concrete at rotating principal axes. *J. Eng. Mech. ASCE* **128**(1), 24–33 (2002)
- F.C. Caner, Z.P. Bažant, Microplane model M7 for plain concrete. I: formulation. *J. Eng. Mech.* **139**(12), 1714–1723 (2013a)
- F.C. Caner, Z.P. Bažant, Microplane model M7 for plain concrete. II: calibration and verification. *J. Eng. Mech.* **139**(12), 1724–1735 (2013b)
- F.C. Caner, Z.P. Bažant, C.G. Hoover, A.M. Waas, K.W. Shahwan, Microplane model for fracturing damage of triaxially braided fiber-polymer composites. *J. Eng. Mater. Tech. Trans. ASME* **133**(2) 021024-1–021024-12 (2011)
- F.C. Caner, Z. Guo, B. Moran, Z.P. Bažant, I. Carol, Hyperelastic anisotropic microplane constitutive model for annulus fibrosus. *J. Biomech. Eng. Trans. ASME* **129**(5), 632–641 (2007)
- J. Červenka, Z. Bažant, M. Wierer, Equivalent localization element for crack band approach to mesh-sensitivity in microplane model. *Int. J. Numer. Methods Eng.* **62**(5), 700–726 (2005)
- K.-T. Chang, S. Sture, Microplane modeling of sand behavior under non-proportional loading. *Comput. Geotech.* **33**(3), 177–187 (2006)
- X. Chen, Z.P. Bažant, Microplane damage model for jointed rock masses. *Int. J. Numer. Anal. Methods Geomech.* **38**(14), 1431–1452 (2014)
- G. Cusatis, A. Beghini, Z.P. Bažant, Spectral stiffness microplane model for quasibrittle composite laminates – part I: theory. *J. Appl. Mech. Trans. ASME* **75**(2) 021009-1–021009-9 (2008)
- G. Cusatis, A. Mencarelli, D. Pelessone, J. Baylot, Lattice discrete particle model (LDPM) for failure behavior of concrete. II: calibration and validation. *Cem. Concr. Compos.* **33**(9), 891–905 (2011)
- G. di Luzio, A symmetric over-nonlocal microplane model {M4} for fracture in concrete. *Int. J. Solids Struct.* **44**(13), 4418–4441 (2007)

- G. Di Luzio, Z. Bažant, Spectral analysis of localization in nonlocal and over-nonlocal materials with softening plasticity or damage. *Int. J. Solids Struct.* **42**(23), 6071–6100 (2005)
- C.G. Hoover, Z.P. Bazant, J. Vorel, R. Wendner, M.H. Hubler, Comprehensive concrete fracture tests: description and results. *Eng. Fract. Mech.* **114**, 92–103 (2013)
- K. Kirane, Z. Bažant, Microplane damage model for fatigue of quasibrittle materials: sub-critical crack growth, lifetime and residual strength. *Int. J. Fatigue* **70**, 93–105 (2015)
- K. Kirane, M. Salviato, Z.P. Bažant, Microplane-triad model for elastic and fracturing behavior of woven composites. *J. Appl. Mech. Trans. ASME***83**(4) 041006-1–041006-14 (2016)
- M.B. Nooru-Mohamed, Mixed-mode fracture of concrete: an experimental approach. Ph.D. thesis, Civil Engineering and Geosciences, Universiteitsdrukkerij, Delft University of Technology, Delft, The Netherlands. (1992, 5)
- J. Ožbolt, Z. Bažant, Microplane model for cyclic triaxial behavior of concrete. *J. Eng. Mech. ASCE* **118**(7), 1365–1386 (1992)
- P. Prat, Z. Bažant, Microplane model for triaxial deformation of saturated cohesive soils. *J. Geotech. Eng. ASCE* **117**(6), 891–912 (1991)
- Simulia Corporation, Abaqus User Subroutines Reference Guide. Dassault Systèmes (2014)
- R. Wendner, J. Vorel, J. Smith, C.G. Hoover, Z.P. Bazant, G. Cusatis, Characterization of concrete failure behavior: a comprehensive experimental database for the calibration and validation of concrete models. *Mater. Struct.* **48**(11), 3603–3626 (2015)



Modeling Temperature-Driven Ductile-to-Brittle Transition Fracture in Ferritic Steels

31

Babür Deliktaş, Ismail Cem Turtuk, and George Z. Voyiadjis

Contents

Introduction	1100
Theoretical Background	1103
Ductile Fracture: Gurson Porous Plasticity	1103
Brittle Fracture Model: Continuum Damage Mechanics	1106
Application to Small Punch Testing	1110
Small Punch Experimental Setup	1110
Numerical Simulations of SPT	1111
Effects of Specimen Geometry on Fracture Response of SPT	1114
Discussion of Results and Further Research	1118
References	1119

Abstract

The most catastrophic brittle failure in ferritic steels is observed as their tendency of losing almost all of their toughness when the temperature drops below their ductile-to-brittle transition (DBT) temperature. There have been put large efforts

B. Deliktaş (✉)

Faculty of Engineering-Architecture, Department of Civil Engineering, Uludag University, Bursa, Turkey

e-mail: bdeliktas@uludag.edu.tr

I. Cem Turtuk

Mechanical Design Department, Meteksan Defence, Ankara, Turkey

Department of Civil Engineering, Uludag Univeristy, Bursa, Turkey

e-mail: ismail.turtuk@gmail.com

G. Z. Voyiadjis

Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge, LA, USA

e-mail: cgzv1@lsu.edu

in experimental and theoretical studies to clarify the controlling mechanism of this transition; however, it still remains unclear how to model accurately the coupled ductile/brittle fracture behavior of ferritic steels in the region of ductile-to-brittle transition.

Therefore, in this study, an important attempt is made to model coupled ductile/brittle fracture by means of blended micro-void and micro-cracks. To this end, a thermomechanical finite strain-coupled plasticity and continuum damage mechanics models which incorporate the blended effects of micro-heterogeneities in the form of micro-cracks and micro-voids are proposed.

In order to determine the proposed model material constant, a set of finite element model, where the proposed unified framework, which characterizes ductile-to-brittle fracture behavior of ferritic steels, is implemented as a VUMAT, is performed by modeling the benchmark experiment given in the experimental research published by Turba et al., then, using these models as a departure point, the fracture response of the small punch fracture testing is investigated numerically at 22°C and −196°C and at which the fracture is characterized as ductile and brittle, respectively.

Keywords

Ductile-Brittle transition · Porous plasticity · Damage mechanics · Small punch test · Ferritic steels

Introduction

Large-scale ferritic steel structures are widely used in advanced engineering applications such as pressure vessels, fusion reactor structures, construction of nuclear reactors, line pipes used for gas and oil transportation, and welded steel ships. It is well known that during the low-temperature climate operation, the catastrophic breakdown of structure and machine elements often occurred. As noted by Sutar et al. (2014) between 1942 and 1952, around 250 large welded steel ships were lost due to catastrophic brittle failure. Although the most of steel structure is normally capable of sustaining great loads and capable of ductility above certain degree of temperature, they become so brittle and have tendency to lose almost all of their toughness when the temperature drops below their ductile-to-brittle transition temperature (Baloso et al. 2017). Depending on loading rate, ambient temperature, as well as triaxiality, the mode of fracture in metals switches from ductile to brittle or vice versa, which is called as ductile-to-brittle transition phenomenon. The ductile-to-brittle transition temperature (DBTT) is a phenomenon that is widely observed in metals especially in severe steel structures in severe weather (Renevey et al. 1996; Tanguy et al. 2007; Hütter 2013b).

In metals, the fracture mode in the upper-temperature range (toward 22°C) is ductile with micro-voids and dimples. At the lowest-temperature range (toward −196°C), the fracture is almost completely brittle, while plasticity governed ductile failure is observed at temperatures as low as −158°C. In that case, the

plasticity properties fall, and the character of fracture changes from ductile to brittle (Hütter 2013b). In the transition region, both ductile and brittle fracture, consisting of transgranular cleavage marks, exist. The amount of ductile crack growth decreases gradually with temperature decrease. At lower temperature, brittle failure is linked with the inter- or intragranular cleavage with nucleation, growth, and coalescence of micro-cracks and is identified as dominant unstable failure mechanisms, while at the higher temperature ductile failure is characterized by the nucleation, growth, and coalescence of micro-voids leading to rupture (Anderson 2004; Chakraborty and Biner 2013).

The controlling mechanism of this transition still remains unclear despite large efforts made in experimental and theoretical investigation. Thus understanding the triggering mechanism of such kind of transition is of great importance. Two common failure mechanisms, namely, ductile and brittle fracture in metallic materials, should be characterized well in the constitutive model for better understanding the triggering mechanism of such kind of transition.

There have been significant efforts put forward to develop unified models that take into account the different failure mechanisms that are active at different temperature regime as well as at transition zone (see, e.g., Shterenlikht 2003; Hutter et al. 2014; Needleman and Tvergaard 2000; Batra and Lear 2004; Xia and Fong Shih 1996; Chakraborty and Biner 2013; McAuliffe and Waisman 2015; Turtuk and Deliktas 2016; Soyarslan et al. 2016).

In those approaches, ductile failure mechanism is modeled by the strip line model of Freund and Lee (1990) and Rousselier (1987)'s model, or mostly by Gurson-Tvergaard-Needleman model (GTN-model Needleman and Tvergaard 2000), whereas cleavage failure in metallic materials has been described primarily by two models. The deterministic model by Ritchie-Knott-Rice (Ritchie et al. 1973) relies on a critical stress over the critical distance principle. In other words, brittle fracture occurs once the principal stresses averaged over a region within a characteristic length exceed a temperature- and rate-independent threshold, while it may or may not be accompanied by plastic flow. Drawback of these models is that the experimentally observed scatter caused by cleavage in fracture toughness cannot be captured. Curry and Knott (1979) modified this theory by incorporating a critical volume ahead of the crack tip to describe the scatter in fracture toughness due to cleavage failure.

The statistical model proposed by Beremin Research Group (see, e.g., Beremin 1983 and Mudry 1987), on the other hand, incorporates Weibull distribution (Weibull 1953) to describe fracture stress in terms of failure probability at lower and transition temperatures. Statistical nature of this model helps one to capture the scatter in fracture toughness; however, it fails to model accurately stable crack growth near transition regime. Turtuk and Deliktas (2016) proposed a new approach specifically to model temperature-driven ductile brittle transition regime to introduce internal softening associated with micro-crack initiation and propagation due to cleavage mechanisms.

Many researchers proposed approaches that use a constitutive model such as the GTN model which accounts for the damage-induced nonlocality

(see, e.g., Bergheau et al. 2014; Zhang et al. 2000; Rivalin et al. 2001a, b; Pardoen and Hutchinson 2000). Linse et al. (2012) has made nonlocal modification on the GTN model by employing an implicit gradient-enriched formulation where the volumetric plastic strain is introduced as nonlocal, damage-driving state variable. Ramaswamy and Aravas (1998) has been discussed and implemented the gradient-based extension of the Gurson model.

There are only few studies in the literature where the particular cleavage softening is taken into account. Samal et al. (2008) modified Rousselier model by considering a nonlocal Rousselier model (Samal et al. 2008) in conjunction with Beremin model (Beremin 1983). Chakraborty and Biner (2013) proposed a unified model that incorporates both ductile damage and cleavage failure mechanisms through temperature- and failure probability-dependent parameters. The flow strength of the bulk material is varied to obtain the temperature-dependent bulk material behavior. It is assumed that without cleavage, the cohesive law follows a traction-separation behavior of ductile damage as described in Scheider and Brocks (2003). Modeling cleavage failure by a cohesive zone is a heuristic approach which combines two relevant features. Firstly, the softening initiates under pure mode-I when the maximum principal stress in the ligament reaches a critical level, the so-called cohesive strength, which is a common assumption. In addition, the work of cohesive separation can be interpreted as the work required to drive the micro-crack which initiated at a broken second-phase particle into the neighboring grain and through the next grain boundary (Kabir et al. 2007; Hardenacke et al. 2012)

In the work of Hutter (2013a), a unified model is proposed by means of a cohesive zone in addition to the ductile material degradation, where the crack propagation is simulated by incorporating the softening associated with cleavage initiation in the ductile-brittle transition region. He introduced two length scales to investigate the material behavior. On the macroscopic scale, the ductile mechanism is modeled with a nonlocal modification of the Gurson-Tvergaard-Needleman model (GTN model), whereas the discrete voids are resolved in the microscopic model. Soyarslan et al. (2016) proposed a model for the ductile-brittle transition at the macroscale where two length scales are incorporated into the constitutive relations. The ductile fracture length scale is based on the average inclusion distance and associated with the nonlocal evolution equation for the porosity. The brittle fracture length scale is based on the average grain size and associated with the material region at which the maximum principal stress is averaged out.

However, these approaches suffer from the mentioned inherent weaknesses pertaining to brittle fracture models. An important attempt to model ductile fracture by means of blended micro-void and micro-cracks is given in literature by Chaboche et al. (2006). In their work, micro-void- and micro-crack-driven separate damage variables have been defined in order to model ductile fracture at isothermal conditions, meaning that no thermal coupling exists in their work. Recently Turtuk and Deliktaş (2016) published a work that presents present thermomechanical finite strain-coupled plasticity and continuum damage mechanics models which incorporate the blended effects of micro-heterogeneities in the form of micro-cracks and micro-voids. They applied presented models in small punch fracture testing in order to simulate temperature-driven ductile-to-brittle transition fracture of P91

steel. While void growth-driven damage is represented by Gurson plasticity, micro-crack-driven damage is formulated with continuum damage mechanics framework. The presented work also takes up blended modeling approach and extends further, incorporating thermal coupling, Gurson's porous plasticity, brittle fracture modeling, and application to ductile-to-brittle transition fracture. The developed unified framework, which characterizes ductile, brittle, and ductile-to-brittle transition fractures, was implemented as a VUMAT in ABAQUS to describe fracture response. Taking the experimental research of notched specimens published by Turba et.al, the finite element analyses have been performed to simulate fracture response of the small punch fracture specimens made of P91 steel at 22°C and -196°C and at which the fracture is characterized as ductile and brittle, respectively.

Theoretical Background

Theoretical framework presented here based on the thermomechanical finite strain-coupled plasticity and continuum damage mechanics models which incorporate the blended effects of micro-heterogeneities in the form of micro-cracks and micro-voids. While void growth-driven damage is represented by Gurson plasticity, micro-crack-driven damage is formulated with continuum damage mechanics framework. The presented work also takes up blended modeling approach and extends further, incorporating thermal coupling, Gurson's porous plasticity, brittle fracture modeling, and application to ductile-to-brittle transition fracture. The developed unified framework, which characterizes ductile, brittle and ductile to brittle transition

Ductile Fracture: Gurson Porous Plasticity

The modeling of fracture in ductile metals due to damage is often based on the micromechanical model of Gurson (1977) for the growth of a single void in an ideal elastoplastic matrix. This model devises a hydrostatic stress-dependent yield potential derived using homogenization over void rigid plastic matrix and limit analysis. This potential is later modified by Tvergaard and Needleman, by the introduction of void shape effects as well as acceleration in the void growth during void coalescence, to be named as Gurson-Tvergaard-Needleman porous plasticity model (Tvergaard and Needleman 1984), along with other contributors, e.g., Tvergaard (1981, 1982a,b), Needleman and Tvergaard (1998), Nahshon and Hutchinson (2008), Nahshon and Xue (2009), Wen et al. (2005), and Malcher et al. (2012), also with extensions to nonlocal formulation at hyper-elastic setting as in Hakansson et al. (2006).

$$\phi^p = \left[\frac{\bar{\sigma}_{eq}}{\sigma_y} \right]^2 + 2q_1 f^* \cosh \left(\frac{q_2}{2\sigma_y} \text{tr} [\bar{\sigma}] \right) - \left[1 + [q_1 f^*]^2 \right] = 0. \quad (1)$$

$$\bar{\sigma}_{eq} = \sqrt{\frac{3}{2} [\text{dev} \bar{\sigma} : \text{dev} \bar{\sigma}]}$$

where f is the void volume fraction that evolves with plastic strains, σ_y is the current yield stress of material, and $\tilde{\sigma}$ is the so-called *effective stress*, which is defined within continuum damage mechanics formalism and introduced in detail in the following section. Note that the proposed coupling with continuum damage mechanics formalism introduces $\text{dev}[\tilde{\sigma}]$ as the *deviatoric part of the effective stress tensor at rotationally neutralized configuration* in the yield potential ϕ^p , onto which both void growth and micro-crack-driven damage mechanisms are reflected in order to account for softening effects of ductile and brittle failure phenomena. Since the aim is to model combined effect of micro-voids and micro-cracks on fracture, the effective stress-based plastic potential can be considered as a Gurson's plastic matrix with built-in micro-cracks. A similar usage of effective stress in the plastic dissipation potential has been introduced by Chaboche et al. (2006), by using two damage state variables (one for pressure-dependent plastic flow response and the other for generalized damage situations) and focusing only on ductile fracture with isotropic damage. Further details on effective stress concept and micro-crack-driven damage are given in the following section.

Evolution of the plastic flow is assumed to follow normality rule and is calculated from:

$$\dot{\epsilon}^p = \lambda \partial_{\tilde{\sigma}} \chi = \lambda \partial_{\tilde{\sigma}} \phi^p \quad (2)$$

where λ is the plastic multiplier, χ is the dissipation potential function which is defined as $\chi = \phi^p(\tilde{\sigma}, \mathbf{D}) + \phi^d(\mathbf{Y}, \mathbf{D})$, with ϕ^d being the damage dissipation potential representing micro-crack-driven damage contribution linked to brittle fracture. Note that the micro-crack-driven damage is coupled to plastic flow upon λ , considering the plastic flow existence in ductile-to-brittle transition state. Further elaboration of damage potential and its parameters are given in subsection on continuum damage mechanics. Tvergaard and Needleman's criteria for void growth and coalescence have been defined by:

$$f^*(f) = \begin{cases} f & , f \leq f_c \\ f_c + [f_u^* - f_c] \frac{[f - f_c]}{[f_F - f_c]} & , f > f_c \end{cases} \quad (3)$$

where $f_u = 1/q_1$, f_c is the threshold value of void volume ratio at which void growth and coalescence start to accelerate and f_F is the critical value of void volume ratio at which the load-carrying capacity of the material is completely lost. The evolution equation for void volume ratio f is calculated using principle of conservation of mass and incompressibility of plastic flow. It is composed of two separate phases, namely, nucleation and growth under fully developed plastic flow, meaning that evolution of void volume fraction can mathematically be defined by:

$$\dot{f} = \dot{f}^{\text{nuc}} + \dot{f}^{\text{gr}} \quad (4)$$

The void evolution due to nucleation is statistically dependent and is defined as in Chu and Needleman (1980):

$$\dot{f}^{\text{nuc}} = A_N \tilde{\varepsilon}_{\text{eq}}^p; \quad A_N = A_N(\tilde{\varepsilon}_{\text{eq}}^p) = \frac{f_N}{S_N \sqrt{2\pi}} \exp\left(-\frac{[\tilde{\varepsilon}_{\text{eq}}^p - \tilde{\varepsilon}_N^p]^2}{2[S_N]^2}\right) \quad (5)$$

where $\tilde{\varepsilon}_N^p$ and S_N denote the mean equivalent plastic strain at the nucleation and its standard deviation. Employing the principle of plastic work equivalence, the equivalent effective plastic strain rate reads as follows.

$$\tilde{\varepsilon}_{\text{eq}}^p = \frac{\bar{\boldsymbol{\sigma}} : \dot{\boldsymbol{\varepsilon}}^p}{[1 - f]\sigma_y} \quad (6)$$

After Nahshon and Hutchinson (2008), \dot{f}^{gr} has further been decomposed into void growth under shear and under normal stresses such that $\dot{f}_{\text{normal}}^{\text{gr}}$ controls void growth under hydrostatic stresses and $\dot{f}_{\text{shear}}^{\text{gr}}$ enhances behavior of softening due to void growth when material is under shear stresses.

$$\begin{aligned} \dot{f}^{\text{gr}} &= \dot{f}_{\text{normal}}^{\text{gr}} + \dot{f}_{\text{shear}}^{\text{gr}} \\ \dot{f}_{\text{normal}}^{\text{gr}} &= [1 - f] \text{tr} \left[\dot{\boldsymbol{\varepsilon}}^p \right] \\ \dot{f}_{\text{shear}}^{\text{gr}} &= k_w f \frac{w}{\bar{\boldsymbol{\sigma}}_{\text{eq}}} \text{dev} [\bar{\boldsymbol{\sigma}}] : \dot{\boldsymbol{\varepsilon}}^p \end{aligned} \quad (7)$$

where k_w is a material parameter governing shear-related growth and is suggested in Nahshon and Hutchinson (2008) to be within $0 \leq k_w \leq 3$ and w is a function of deviatoric stress within $0 \leq w \leq 1$ that is responsible for distinguishing the axisymmetric stress states from generalized plane strain and is calculated as:

$$w = \hat{w}(\text{dev} [\bar{\boldsymbol{\sigma}}]) = - \left[\frac{27 \bar{J}_3}{2 \bar{\boldsymbol{\sigma}}_{\text{eq}}^3} \right]^2 \quad (8)$$

in which \bar{J}_3 is the third deviatoric invariant of the stress tensor. Note that, although a generalized shear-driven evolution is included within the formalism of continuum damage mechanics, within the formulation of the present work, Nahshon and Hutchinson shear correction is required to describe damage for ductile fracture, while CDM formulation is taken to model brittle damage evolution. Finally, the balance of energy is used to compute the time rate of change of temperature using the plastic work equation:

$$\rho c_p \dot{\theta} = \zeta \bar{\boldsymbol{\sigma}} : \dot{\boldsymbol{\varepsilon}}^p \quad (9)$$

where ζ denotes the so-called Taylor-Quinney coefficient (Taylor and Quinney 1934) that controls the amount of plastic work converted to heat during plastic flow, in which ρ and c_p are the density and the heat capacity of the material, respectively. Extension of the plastic hardening model for strain rate and temperature dependency has been performed considering Johnson-Cook-type multiplicative decomposition of the material yield stress on strain hardening, strain rate dependency, and thermal softening. However, the original Johnson-Cook hardening law has been replaced with Hockett-Sherby hardening equation (Hockett and Sherby 1975), for better representation of the post-yielding and pre-peak response in small punch testing simulations of P91. That is:

$$\begin{aligned}\sigma_y(\tilde{\varepsilon}_{\text{eq}}^p, \dot{\tilde{\varepsilon}}_{\text{eq}}^p, \theta) &= h_y(\tilde{\varepsilon}_{\text{eq}}^p) r_y(\dot{\tilde{\varepsilon}}_{\text{eq}}^p) t_y(\theta) \\ h_y(\tilde{\varepsilon}_{\text{eq}}^p) &= h \tilde{\varepsilon}_{\text{eq}}^p + b - [b - \sigma_{y0}] \exp(-m [\tilde{\varepsilon}_{\text{eq}}^p]^n) \\ r_y(\dot{\tilde{\varepsilon}}_{\text{eq}}^p) &= 1 + C \log \left(\frac{\dot{\tilde{\varepsilon}}_{\text{eq}}^p}{\dot{\varepsilon}_0^p} \right) \\ t_y(\theta) &= 1 - \Omega^r\end{aligned}\tag{10}$$

where σ_{y0} , b , C , n , m and r are material parameters. The reference strain rate has been denoted with $\dot{\varepsilon}_0^p$, and Ω is defined as $\Omega := [\theta - \theta_0]/[\theta_M - \theta_0]$ where θ_0 and θ_M represent the reference and melting temperatures, respectively. Here, only dependency of the yield stress on strain rate and temperature is considered within Johnson-Cook formalism, i.e., employed formulation is not viscoplastic. Note that theoretical basis presented here and numerical implementation performed cover also rate effects. However, it is excluded within the scope of this work, and only temperature-driven effects are investigated.

Brittle Fracture Model: Continuum Damage Mechanics

Concept of damaged material has initially been introduced by Kachanov (1958) and Rabotnov (1969), with an attempt to describe creep behavior of materials. Following these initial proposals, the concept of damage mechanics has been developed through works of Chaboche (1977), Lemaitre (1985), Germain et al. (1983), Murakami and Ohno (1980), Chow and Wang (1987a,b), and Krajcinovic and Fonseka (1981), within consistency of continuum thermodynamics. An exhaustive literature cite is not aimed here; however further literature is given in Besson et al. (2010) and Skrzypek et al. (2008).

For a general case, damage is physically defined as the ratio of the surface area of micro-cracks or cavities situated on the plane passing through the cross section of a representative volume element (RVE) to the total surface area of the RVE. Damage is assumed to act over an RVE such that $D = 0$ describes undamaged material,

while $D = 1$ represents fully damaged material, at which the load-carrying capacity of RVE has totally been lost. In phenomenological approach to damage, also called as continuum damage mechanics, a damage state variable affects the material stress state, with the introduction of the so-called *effective* stress tensor. Definition of the effective stress tensor is either based on the strain equivalence principle introduced by Chaboche (1977) or by the energy equivalence principle developed by Cordebois and Sidoroff (1982). In strain equivalence principle, effective stress is defined as the stress tensor, when applied to a fictitious undamaged material generates the same strain tensor obtained when net (or homogenized) stress tensor is applied to a damaged material. In energy equivalence principle, the equivalence is formulated over elastic strain energies of damaged and undamaged materials. For the case of isotropic damage, effective stress is defined as $\tilde{\sigma} = \tilde{\sigma}/(1 - d)$, where d is a scalar damage variable and σ is the net stress tensor. For the tensorial damage variable, also employed in the present work, several definitions are given in literature (Skrzypiek et al. 2008).

The effective stress tensor employed in the present work is based on the volumetric-deviatoric split and symmetrization of effective stress tensor given by Cordebois and Sidoroff (1982) and Chow and Wang (1987a,b) and as:

$$\tilde{\sigma} = \frac{\text{tr}[\tilde{\sigma}]}{3 - \text{tr}[\mathbf{D}]} \mathbf{1} + \left[[\mathbf{1} - \mathbf{D}]^{-\frac{1}{2}} \text{dev}[\tilde{\sigma}] [\mathbf{1} - \mathbf{D}]^{-\frac{1}{2}} \right] : \mathbb{P} \quad (11)$$

$$\tilde{\sigma} = \frac{\text{tr}[\tilde{\sigma}]}{3 - \text{tr}[\mathbf{D}]} \mathbf{1} + \left[[\mathbf{1} - \mathbf{D}]^{-\frac{1}{2}} \text{dev}[\tilde{\sigma}] [\mathbf{1} - \mathbf{D}]^{-\frac{1}{2}} \right] : \mathbb{P} \quad (12)$$

in which “tr” and “dev” denote trace and deviator of the tensorial quantity and \mathbb{P} is the 4th order deviatoric projection tensor and \mathbf{D} is the 2nd order tensorial damage variable representing the micro-crack-driven damage evolution linked to brittle fracture. The coupling and dependency between micro-crack-driven brittle damage evolution and micro-void-driven ductile damage is achieved by assuming the normality rule for damage evolution, using plastic multiplier λ and potential $\chi = \phi^p(\tilde{\sigma}, \mathbf{D}) + \phi^d(\mathbf{Y}, \mathbf{D})$ introduced previously, with a specific choice of damage potential ϕ^d given in Hayakawa and Murakami (1997) and Abu Al-Rub and Voyiadjis (2003):

$$\begin{aligned} \dot{\mathbf{D}} &= \lambda \partial_{\mathbf{Y}} \chi = \lambda \partial_{\mathbf{Y}} \phi^d \\ \phi^d &= \frac{1}{S^m(m+1)} \sqrt{[\mathbf{1} - \mathbf{D}]^{-1} : [\mathbf{1} - \mathbf{D}]^{-1}} \sqrt{[\mathbf{Y} : \mathbf{Y}]^{m+1}} \end{aligned} \quad (13)$$

where \mathbf{Y} is stress-like variable, controlling brittle damage initiation and evolution, and is conjugate to \mathbf{D} . Note that, the specific choice for ϕ^d is in terms of the invariants of stress-like damage conjugate variable \mathbf{Y} so that the brittle damage formulation is analogous to ductile damage part with ϕ^p and f . In a thermodynam-

ically consistent framework, one can obtain the expression for damage conjugate variable named as strain energy release rate \mathbf{Y} , by taking its derivative with respect to the free energy function defined, i.e., $\mathbf{Y} = -\partial_{\mathbf{D}}\psi$. Instead, for the hypo-elastic plasticity framework upon which this work is based, we propose two different laws to calculate \mathbf{Y} and $\dot{\mathbf{D}}$. Looking from brittle fracture point of view and based on experimental evidence that brittle fracture is linked to principal stresses, principal stress-driven Leckie-Hayhurst-type and plasticity-driven Lemaitre-type expressions for \mathbf{Y} and $\dot{\mathbf{D}}$ have been proposed first time specifically to model temperature-driven ductile-to-brittle transition. These proposed models are introduced and discussed next.

Modified Leckie-Hayhurst Form

Initially developed as a creep-rupture criterion (Hayhurst and Leckie 1973; Leckie and Hayhurst 1977) in terms of stress invariants and within the context of isotropy, the so-called Leckie-Hayhurst form is generalized as a function of $\bar{\boldsymbol{\sigma}}$ as:

$$\varphi(\bar{\boldsymbol{\sigma}}) = a\bar{\sigma}_i + bJ_1(\bar{\boldsymbol{\sigma}}) + [1 - a - b]J_2(\bar{\boldsymbol{\sigma}}) \quad (14)$$

where $\bar{\sigma}_i$ is the maximum principal component of effective stress tensor at rotationally neutralized configuration. Proposing such a form for strain energy release rate has two advantages: (a) there is direct dependency to stresses and therefore direct link to brittle fracture modeling over principal stresses and (b) straightforward numerical implementation. Furthermore, the proposed form is analogous to quantitative expression of Ritchie-Knott-Rice cleavage criterion, which requires maximum principal stress exceed a specific temperature-independent limit over a specific distance. Thus, Leckie-Hayhurst-type form has been modified to capture principal stress effects such that $Y_i = \bar{\sigma}_i$. Since the form in general has been derived considering isotropy condition (over a scalar damage variable), the extension considering second-order damage tensor \mathbf{D} has been proposed in terms of spectral representation of \mathbf{Y} :

$$\mathbf{Y} = \sum_{i=1}^3 \langle Y_i \rangle \mathbf{n}_i \otimes \mathbf{n}_i \quad (15)$$

$$\langle Y_i \rangle = \max(0, Y_i - Y_i^0)$$

where Y_i^0 is the threshold value controlling the micro-crack damage initiation and evolution and \mathbf{n}_i the principal directions of effective stress tensor $\bar{\boldsymbol{\sigma}}$. This modified form can also be considered as 2nd-order anisotropic Ritchie-Knott-Rice brittle fracture criterion which is able to model internal softening associated with micro-crack initiation, propagation, and thus cleavage. However, note that this proposed form for \mathbf{Y} has no more the meaning of strain energy release rate, as it is in a thermodynamically consistent framework.

For the calibration of the material parameter Y_i^0 , a mixed experimental-numerical study needs to be performed. Here, we propose a similar test setup given in Ritchie et al. (1973) and Curry and Knott (1976). In Ritchie et al. (1973) and Curry and Knott (1976), a four-point bending test has been set up, and fracture toughness of the pre-notched specimens have been measured. A finite element study of the stress field around the notch then revealed the critical fracture stress of the specimen. Since the principal stress dependency of the proposed Leckie-Hayhurst form is similar to that of Ritchie-Knott-Rice fracture law, a parameter identification setup similar to Ritchie et al. (1973) and Curry and Knott (1976), which measures fracture response in terms of fracture toughness or force-displacement curves then correlates it to principal stresses numerically, is sufficient for the determination of Y_i^0 . Note that, the experimental portion of the correlations performed within the present work is based on force-displacement response of the small punch fracture experiments given in Turba et al. (2011).

Plasticity-Driven Lemaitre Form

Driven by motivation that micro-void-based dilatational damage f and micro-crack-driven brittle damage coexist in a ductile-to-brittle transition state, the kinetic damage evolution law proposed by Lemaitre et al. (2000) is adopted to describe damage evolution effects to model cleavage. The form proposed by Lemaitre et al. in (2000) is:

$$\dot{\mathbf{D}} = \left[\frac{\bar{Y}}{S} \right]^m |\dot{\tilde{\boldsymbol{\epsilon}}}^p| \quad \text{where} \quad \bar{Y} = \frac{1}{2} \tilde{\boldsymbol{\epsilon}}^e : \mathbb{C} : \tilde{\boldsymbol{\epsilon}}^e \quad (16)$$

where \bar{Y} is the effective elastic energy density, while $|\dot{\tilde{\boldsymbol{\epsilon}}}^p|$ defined plastic strain rate tensor in terms of its principal components, S and m are material parameters to be correlated. $|\dot{\tilde{\boldsymbol{\epsilon}}}^p|$ is formed by rewriting plastic strain rate tensor in spectral representation with its eigenvectors serving a basis and performing absolute value operation. It's to pay attention that principal directions of damage rate tensor coincide with the principal directions of plastic strain rate tensor. The evolution of damage is positive only if a certain threshold of accumulated plastic strain is reached, i.e.,

$$\dot{\mathbf{D}} = \mathbf{0} \quad \text{if} \quad \tilde{\boldsymbol{\epsilon}}_{\text{eq}}^p \leq \tilde{\boldsymbol{\epsilon}}_{\text{eq}}^{p,\text{crit}} \quad (17)$$

The critical value of accumulated equivalent plastic strains acts as a threshold not only for damage evolution but as a limiter to pre-, post-, and in-transition zone fracture response of the material. Therefore, it should be identified along with other damage parameters. Such identification is performed numerically within the scope of the work and is given in preceding sections. A drawback of this form is this: a nonphysical damage evolution is recovered under biaxial tensile stress states. To overcome this issue, Eq. 16 can be modified to take positive part of $\dot{\tilde{\boldsymbol{\epsilon}}}^p$, instead of its absolute value. However, for axisymmetric small punch fracture modeling

and implementation, the original form with absolute value of principal components has been used. The experimental characterization of material parameters S, m , and $\varepsilon_{\text{eq}}^{p, \text{crit}}$ that define this damage evolution law is basically uniaxial test with unloadings and low-cycle fatigue tests. Details of experimental characterization are given in Lemaitre et al. (2000) and Lemaitre and Desmorat (2005).

Application to Small Punch Testing

Small punch testing is a recently developed mechanical testing method that allows the mechanical, fracture, and creep characterization of tiny specimens. It gained widespread importance and popularity since it enables the testing using specimens taken from already existing components that have been exposed to service loads and temperatures. As it is also considered a nondestructive test method, in situ characterization of fracture, creep, mechanical properties, as well as lifetime is possible.

Small Punch Experimental Setup

Small punch testing technique uses miniaturized samples, and its setup has been carried out according to the guidelines of the *Code of Practice for Small Punch Testing for Tensile and Fracture Behavior* (CEN 2006). The testing temperature range was between 22°C and −196°C. In order to achieve the lowest test temperature, the testing equipment has been placed in a vessel filled with liquid nitrogen inside its environmental chamber. The thermocouples with an accuracy of $\mp 2^\circ\text{C}$ were used. The deformation velocity of the samples was kept at constant value of 0.005 mm/s. Specimens were disks of 8 mm diameter and 1 mm thickness, with a predefined circumferential notch. The V-notches were produced by electro-discharge machining, resulting in a notch of diameter 2.5 mm, notch tip radius of 5 microns, and notch depth of 0.5 mm, as proposed in Turba et al. (2011). The schematic of the SP creep testing equipment is presented in Fig. 1.

In this test method, a rigid spherical ended puncher penetrates through a disk specimen at a constant displacement rate. During the experiment, the puncher force F and the midspan central displacement δ are recorded. The importance of small punch testing lies in its versatility: small punch force-displacement curves can be used to estimate mechanical properties such as yield stress, maximum strength, elasticity modulus, fracture toughness, ductile-to-brittle transition temperature, high temperature creep properties, etc. Furthermore, it requires small-sized specimens so that tiny cuts from existing structures could be utilized to investigate long-term in situ mechanical or thermal response of the material. As a result of experiments performed within a wide temperature range, it has been observed that the fracture mode in the upper temperature range (toward 22°C) is ductile with micro-voids and dimples. In the transition region, both ductile and brittle fracture, consisting of transgranular cleavage marks, exist. The amount of ductile crack growth decreases

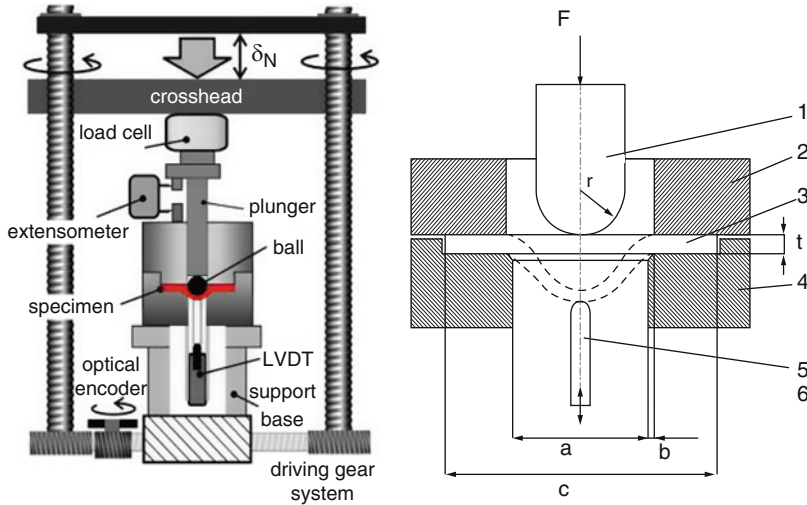


Fig. 1 Small punch test holder (Basbus et al. 2014)

gradually with reduced temperature. At the lower temperature range (toward -196°C), the fracture is almost completely brittle.

The radial symmetry of the notch enables to achieve a plane strain condition for evaluation purposes. Due to significantly small radius of the notch tip, high stress concentrations are achieved, and the crack was found to be initiating from the notch tip, for all the disks tested at different temperatures.

For the ductile case where void initiation and growth are controlled by plastic strain, the crack propagation was found to follow a path in the direction of the notch where the equivalent plastic strain is maximum. The micro-voids nucleating at second phase particles in the vicinity of the notch tip grow and coalescence as the strain level is increased, and finally a continuous crack is formed, leading to the complete fracture of the disk.

At -196°C , the crack propagates in perpendicular direction to maximum principal stress, in accordance with the maximum tensile stress theory by (Erdogan and Sih 1963).

As for the transition, the crack was found to follow an angle in-between ductile and brittle cases, while the angle approaches the one in the brittle case. Ductile, brittle, and transition fractures could be visualized in Figs. 2 and 3.

Numerical Simulations of SPT

The proposed theoretical model has been implemented in ABAQUS as a VUMAT. Numerical implementation has been performed for axisymmetric elements, due to the presence of the rotationally symmetric notch. Implemented model covers

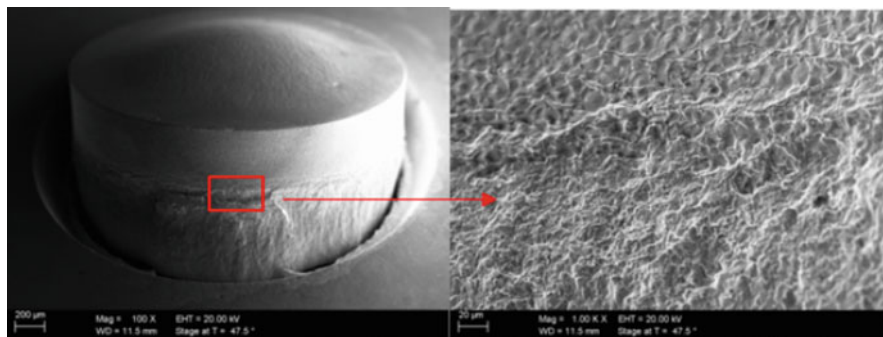


Fig. 2 After-experiment investigation of a notched specimen tested at 22°C



Fig. 3 Three different crack patterns observed at different temperatures (a): ductile fracture at 22°C, (b) brittle fracture at -196°C, (c) temperature-dependent transition fracture pattern

temperature effects on transition fracture as well as Johnson-Cook strain rate dependency and stress triaxiality effects on fracture. Axisymmetric finite elements with reduced integration and enhanced hourglassing controls (CAX4R) have been used. Clamps and puncher have been modeled as rigid instances. Artificial strain energy effects are kept at minimum thanks to sufficiently small element sizes. Small punch experiment has been modeled for 22°C and -196°C. ABAQUS setup can be visualized in Fig. 4.

The set of material parameters used to predict P91 behavior have been listed in Tables 1, 2, 3, and 4.

Simulations with the proposed material model and with stated material parameters have yielded three distinct crack patterns, which are in good correlation both with respect to theoretic predictions and experimental results. If Fig. 5 is inspected, a precise correlation with simulation and numerics, in terms of crack patterns, can easily be observed. Depending on the temperature level θ and the threshold value of the principal strain energy release rate Y_i^0 , direction and pattern of the transition fracture can precisely be predicted. As also inspected from experimental results, transition fracture patterns are mainly dominated by brittle fracture, while plastic strain amount and direction govern how much the crack deviates from principal stress directions. Such experimental and numerical patterns could be visualized in Fig. 5.

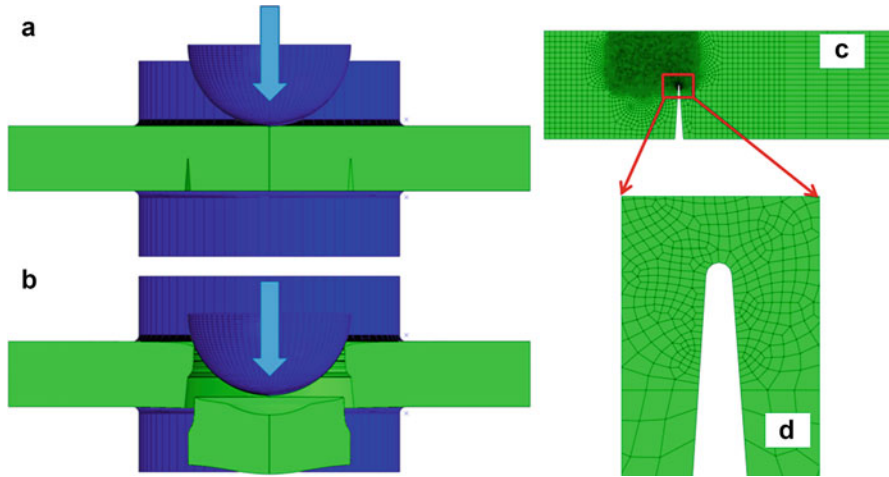


Fig. 4 (a) Initial configuration at the beginning of the simulation. (b) Final configuration at end of the simulation with ductile fracture. (c) Axisymmetric FE mesh of a specimen. (d) Close-up view of the notch tip, illustrating mesh size

Table 1 Thermoelastic properties of P91

E (GPa)	ν	c_p (m ² /[Ks ²])	ζ	ε_0	θ_M (K)
210	0.3	622	0.9	1.3e-5	1717

Table 2 Flow curve data of P91 at 22°C, Johnson-Cook hardening

h (MPa)	σ_{y0} (MPa)	b (MPa)	m	n
500	346	650	25.73	1.01

Table 3 Gurson-Tvergaard-Needleman model parameters for P91

f_N	s_N	ε_N	q_1	q_2	$q_3 = 1/f_U$	f_0	k_w
0.005	0.1	0.03	1.5	1.0	2.25	0.0055	3.0

Table 4 Ductile and brittle failure criteria

f_c	f_F	$D_{i,crit}$	Y_i^0 (MPa)
0.05	0.207	0.30	1500

The force-displacement responses of the implemented models and comparison to experimental results are given in Fig. 6. As it is seen, the numerical response of the proposed models compared to experimental results is in good correlation, as a result of a change in temperature. The peak responses are accurately predicted at all three cases. The maximum pre-failure load for modified Leckie-Hayhurst model is slightly higher compared to that of Lemaitre model. The reason for such a response is natural, since micro-crack-driven damage in Lemaitre-type model is driven by equivalent plastic strain amount, leading to additional softening in load response at higher plastic strains. Still, such a deviation is within the acceptable range. Note that the brittle damage response of modified Leckie-Hayhurst model is linked only to maximum principal strain evolution.

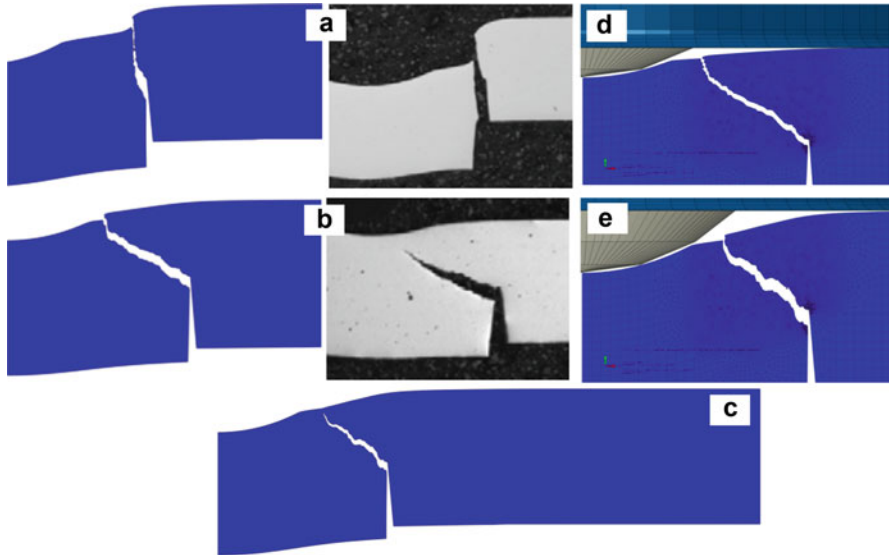


Fig. 5 (a) Numerical prediction and experimental investigation of ductile tearing at 22°C. (b) Numerical prediction and experimental investigation of brittle cleavage at -196°C . (c) A typical mixed-mode fracture pattern, mainly effected by brittle cleavage, while plastic strain direction deviates crack direction away from principal stress directions. (d) Close-up view of pure brittle crack propagation. (e) Close-up view of mixed-mode crack propagation. Note the difference in propagation direction compared to pure brittle fracture

Effects of Specimen Geometry on Fracture Response of SPT

The geometrical parameters of the small punch experiment such as puncher radius and notch depth (Fig. 7) have been varied, and the effects of such geometrical changes on fracture response of the specimen are investigated numerically.

A design of experiments study is performed to find the correlation between geometrical parameters.

The Effects of the Puncher Radius

The puncher radius of the loading units has been varied with the values of 2.25, 2.5, and 2.75 mm, respectively (Fig. 8), to investigate its effects on the fracture response of the notched specimen.

Usually, the radius of the puncher on the loading unit is 2.5 mm. In the case of reducing the radius of the puncher, the contact surface between the puncher and the specimen became smaller. If the puncher radius is reduced to too small value, more stress concentration occur nearby the underneath of the puncher region, and also failure of the specimen is dominated by shear due to punching effect. (See Fig. 9.)

By comparing the experimental results, it is concluded that the puncher radius should be kept between the values of 1.5 and 3.5 mm as its lower and upper limits.

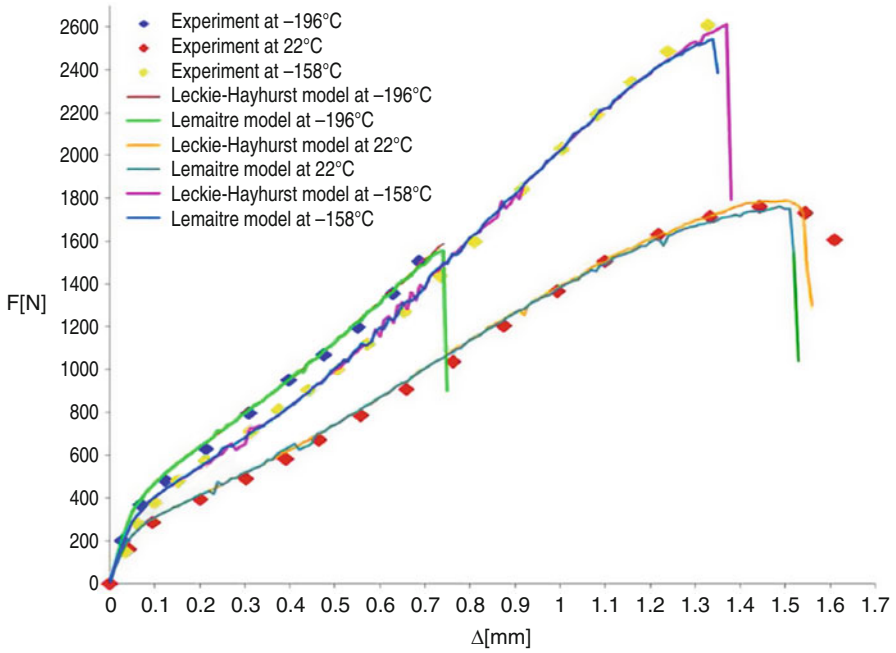


Fig. 6 Numerical simulations with two implemented material models and comparison with experimental results

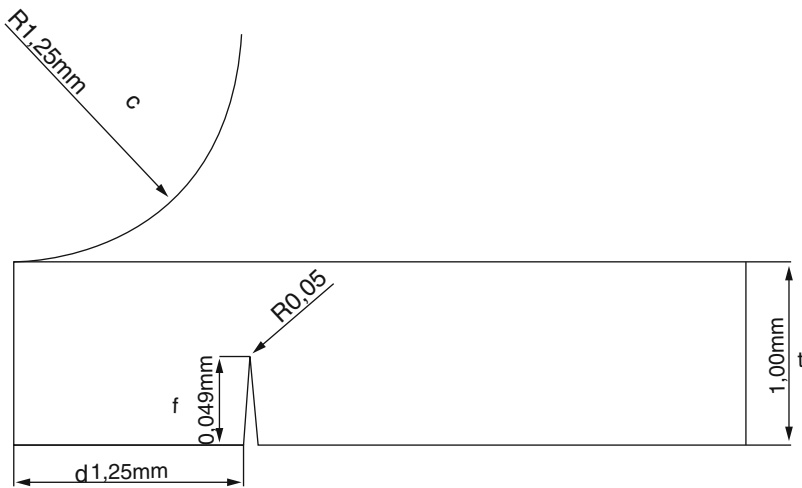


Fig. 7 Geometric parameters of SPT

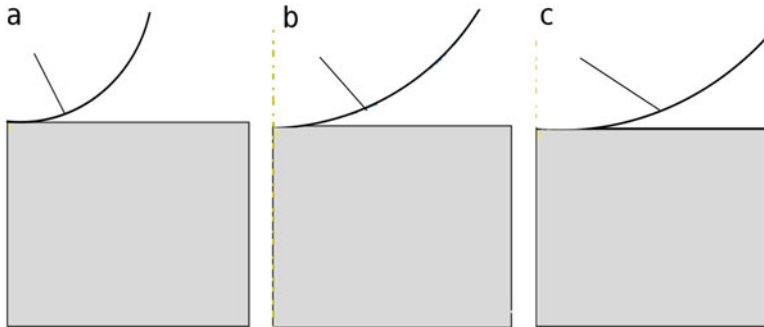


Fig. 8 Various puncher radius

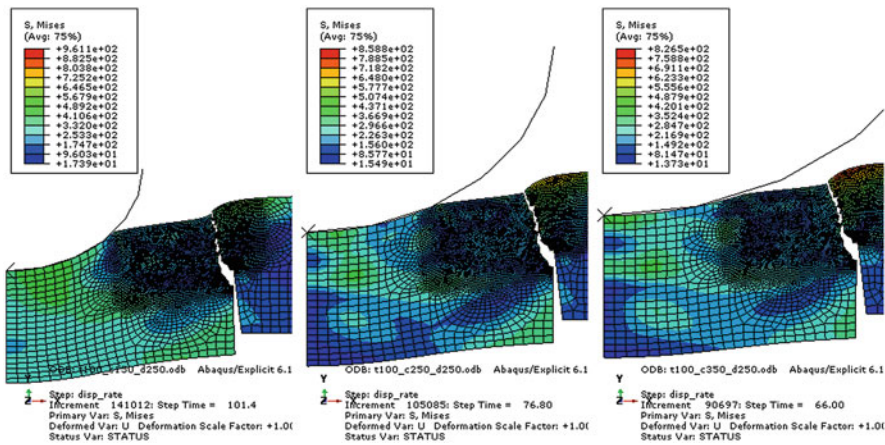


Fig. 9 (a) Numerical prediction of effects of the puncher radius with 2.25, 2.5, and 2.275 mm on the notched specimen

Force-displacement curves are presented in Fig. 10 in order to indicate the puncher radius effect on the fracture response of the specimen.

As it can be seen from the curves in Fig. 10, there is relation between the peak load and the radius of the puncher. As the radius of the puncher increases to a certain critical value, the region of the contacting surfaces between the puncher and the specimen becomes wider and the load reaches its peak value faster than that in the specimen loaded with the puncher having smaller radius. Therefore, this indicates that fracturing of the specimen loaded with puncher having bigger radius occurs much earlier than that of the specimen loaded with puncher having smaller radius (Fig. 9).

The Effects of Crack Depth

Depth of the crack is one of the most important fracture parameters that causes the damage and fracture in the materials. In order to investigate the effect of crack depth

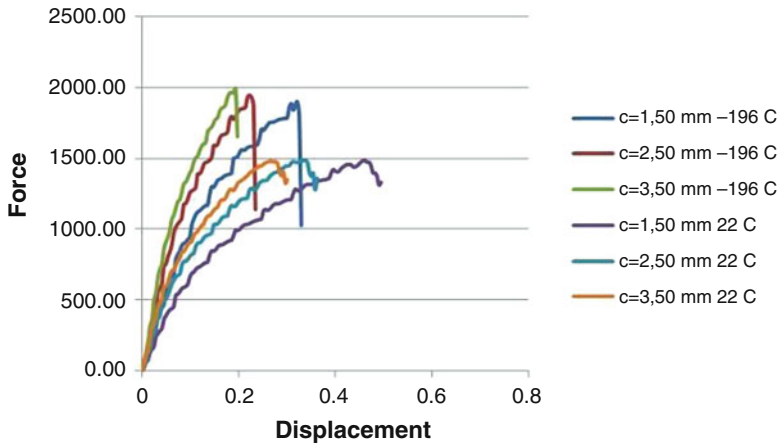


Fig. 10 Prediction of the force-displacement response of the notched specimen for various radius of the puncher

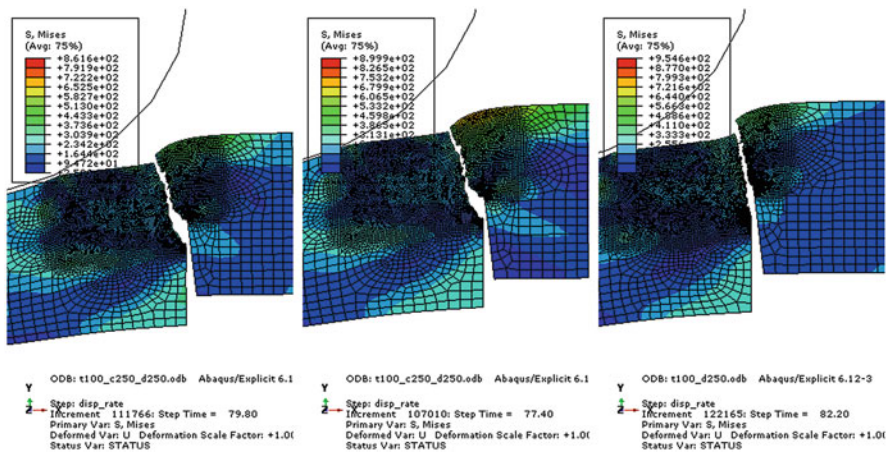


Fig. 11 Stress distribution at various crack depth

on the fracture response of the small punch test, three different crack depths, with length of 0.44, 0.49, and 0.54 mm, are defined in numerical simulation. Figure 11 mimics von Mises stress distributions as the results of the finite element simulation of SPTs at various crack depths.

As the crack depth increases along the direction of the thickness of the specimen, material becomes weaker and is much likely prone to damage and fracture. At the lower temperature -196°C for all crack lengths, the material behaves completely brittle. However, as seen in Fig. 11 for the larger crack length, material shows more ductile behavior than that of the smaller crack length.

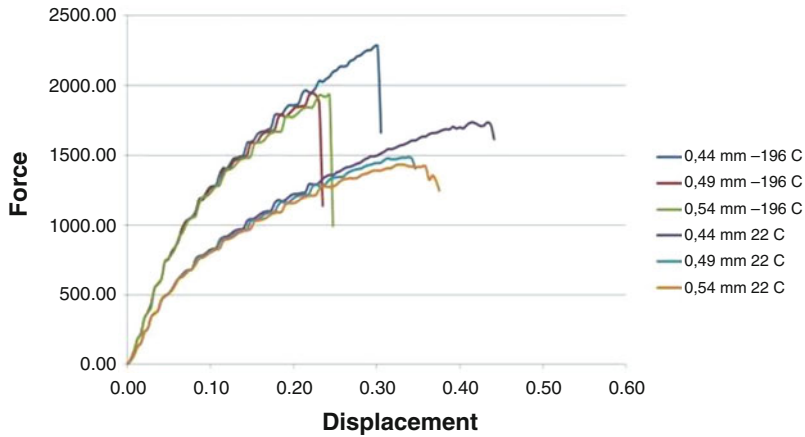


Fig. 12 Prediction of the force-displacement response of the notched specimen for various crack depth

Figure 12 shows the force-displacement curves. It is seen that the peak load of the force-displacement curves exhibit a lowering with increase in the specimen crack length at both temperatures. This implies that for decreased initial crack lengths, the specimen exhibits higher fracture resistance (Fig. 12).

Discussion of Results and Further Research

In the enclosed work, a coupled continuum damage mechanics and porous plasticity model on local scale have been presented, in order to model temperature-dependent ductile-to-brittle transition fracture in metals. Classical effective stress and damage expressions of CDM have been employed to model brittle fracture, while Gurson's porous metal plasticity has been utilized to model ductile fracture. Leckie-Hayhurst criterion has been modified to account for brittle fracture and extended to cover anisotropic material behavior. Constitutive modeling and small punch experiments have been performed for P91 steel. It has been found out that the proposed model precisely and conveniently predicts crack patterns within a relatively wide temperature range from 22°C to -196°C.

The effects of the various geometrical parameters such as puncher radius on fracture response of the SPT specimen is investigated numerically. Depth of the crack is a crucial effect parameter that causes the damage and fracture in the materials. As the crack depth increases along the direction of the thickness of the specimen, material becomes weaker and shows more ductile behavior. On the other hand the puncher radius dominates shear type failure in the specimen due to punching effects.

For the sake of completeness, the proposed model will be extended to nonlocal framework by making use of integral type nonlocal approach, complemented with associated length scales for both brittle and ductile fractures. Incorporation of the statistical effects for the brittle fracture by the stochastic distribution of material properties as well as the sensitivity study for the threshold value of the strain energy release rate Y_i^0 will also be studied.

References

- CEN, Cen workshop agreement: small punch test method for metallic materials. Technical Report CWA 15627 (2006)
- R. Abu Al-Rub, G. Voyiadjis, On the coupling of anisotropic damage and plasticity models for ductile materials. *Int. J. Solids Struct.* **40**, 2611–2643 (2003)
- T. Anderson, *Fracture Mechanics: Fundamentals and Applications*, Taylor & Francis Group, Boca Raton (CRC Press, 2004)
- P. Baloso, R. Madanrao, M. Mohankumar, Determination of the ductile to brittle transition temperature of various metals. *Int. J. Innov. Eng. Res. Technol.* **4**, 1–27 (2017)
- F. Basbus, M. Moreno, A. Caneiro, L. Mogni, Effect of pr-doping on structural, electrical, thermodynamic, and mechanical properties of BaCeO₃ proton conductor. *J. Electrochem. Soc.* **161**(10), 969–976 (2014)
- R. Batra, M. Lear, Simulation of brittle and ductile fracture in an impact loaded prenotched plate. *Int. J. Fract.* **126**, 179–203 (2004)
- F.M. Beremin, A local criterion for cleavage fracture of a nuclear pressure vessel steel. *Met. Trans. A* **14A**, 2277–2287 (1983)
- J.-M. Bergheau, J.-B. Leblond, G. Perrin, A new numerical implementation of a second-gradient model for plastic porous solids, with an application to the simulation of ductile rupture tests. *Comput. Method. Appl. M* **268**, 105–125 (2014)
- J. Besson, G. Cailletaud, J. Chaboche, S. Forest, *Non-linear Mechanics of Materials* (Springer, Berlin, 2010)
- J. Chaboche, Sur l'utilisation des variables d'état interne pour la description de la viscoplasticité cyclique avec endommagement, in *Symposium Franco-Polonais de Rhéologie et Mécanique*, 1977, pp 137–159
- J. Chaboche, M. Boudifa, K. Saanouni, A CDM approach of ductile damage with plastic compressibility. *Int. J. Fract.* **137**, 51–75 (2006)
- P. Chakraborty, B. Biner, Modeling the ductile brittle fracture transition in reactor pressure vessel steels using a cohesive zone model based approach. *Int. Work. Struct. Mater. Innov. Nucl. Syst.* (72):1–10 (2013)
- C. Chow, J. Wang, An anisotropic theory of continuum damage mechanics for ductile fracture. *Eng. Fract. Mech.* **27**, 547–558 (1987a)
- C. Chow, J. Wang, An anisotropic theory of elasticity for continuum damage mechanics. *Int. J. Fract.* **33**, 3–18 (1987b)
- C. Chu, A. Needleman, Void nucleation effects in biaxially stretched sheets. *J. Eng. Mater. Technol.* **102**, 249–256 (1980)
- J.P. Cordebois, F. Sidoroff, Damage induced elastic anisotropy. In: JP. Boehler, (eds), *Mech. Behav. Anisotropic Solids / Comportment Mech. des Sol. Anisotropic*. Springer, Dordrecht https://doi.org/10.1007/978-94-009-6827-1_44 761–774 (1982)
- D. Curry, F. Knott, The relationship between fracture toughness and micro-structure in the cleavage fracture of mild steel. *Mater. Sci.* **10**, 1–6 (1976)
- D. Curry, J.F. Knott, Effect of microstructure on cleavage fracture toughness in mild steel. *Metal Sci.* **13**, 341–345 (1979)

- F. Erdogan, G.C. Sih, On the crack extension in plates under plane loading and transverse shear. *J. Bas. Eng.* **85**, 519–529 (1963)
- L.B. Freund, Y.J. Lee, Observations on high strain rate crack growth based on a strip yield model. *Int. J. Fract.* **42**, 261–276 (1990)
- P. Germain, Q. Nguyen, P. Suquet, Continuum thermodynamics. *J. Appl. Mech.* **50**, 1010–1020 (1983)
- A.L. Gurson, Continuum theory of ductile rupture by void nucleation and growth: Part I- Yield Criteria and Flow Rules for Porous Ductile Media, *J. Eng. Mater. Technol.* **99**, 2–15 (1977)
- P. Hakansson, M. Wallin, M. Ristinmaa, Thermomechanical response of non-local porous material. *Int. J. Plasticity* **22**, 2066–2090 (2006)
- V. Hardenacke, J. Hohe, V. Friedmann, D. Siegele, Enhancement of local approach models for assessment of cleavage fracture considering micromechanical aspects, in *Proceedings of the 19th European Conference on Fracture*, Kazan (72), pp. 49–72 (2012)
- K. Hayakawa, S. Murakami, Thermodynamical modeling of elastic-plastic damage and experimental validation of damage potential. *Int. J. Damage Mech.* **6**, 333–362 (1997)
- D. Hayhurst, F. Leckie, The effect of creep constitutive and damage relationships upon rupture time of solid and circular torsion bar. *J. Mech. Phys. Solids* **21**, 431–446 (1973)
- J. Hockett, O. Sherby, Large strain deformation of polycrystalline metals at low homologous temperatures. *J. Mech. Phys. Solids* **23**(2), 87–98 (1975)
- G. Hutter, Multi-scale simulation of crack propagation in the Ductile-Brittle transition region. PhD. Thesis, Faculty of Mechanical, Process and Energy Engineering, Technische Universität Bergakademie Freiberg (2013a)
- G. Hütter (2013b) Multi-scale simulation of crack propagation in the ductile-brittle transition region. PhD. Thesis, Faculty of Mechanical, Process and Energy Engineering, Technische Universität Bergakademie Freiberg
- G. Hutter, T. Linse, S. Roth, U. Mühlich, M. Kuna, A modeling approach for the complete Ductile-to-Brittle transition region: cohesive zone in combination with a nonlocal gurson-model. *Int. J. Fract.* **185**, 129–153 (2014)
- R. Kabir, A. Cornec, W. Brocks, Simulation of quasi-brittle fracture of lamellar tial using the cohesive model and a stochastic approach. *Comput. Mater. Sci.* **39**, 75–84 (2007)
- L. Kachanov, Time of the rupture process under creep conditions. *IsvAkadNaukSSR* **8**, 26–31 (1958)
- D. Krajcinovic, G. Fonseka, The continuous damage theory of brittle materials, parts 1 and 2. *J. Appl. Mech.* **48**, 809–824 (1981)
- F. Leckie, D. Hayhurst, Constitutive equations for creep rupture. *Acta Metall.* **25**, 1059–1070 (1977)
- J. Lemaitre, A continuous damage mechanics model for ductile fracture. *J. Eng. Mater. Technol.* **107**, 83–89 (1985)
- J. Lemaitre, R. Desmorat, *Engineering Damage Mechanics* (Springer, Berlin/New York, 2005)
- J. Lemaitre, R. Desmorat, M. Sauzay, Anisotropic damage law of evolution. *Eur. J. Mech. Solids* **19**, 187–208 (2000)
- T. Linse, G. Hütter, M. Kuna, Simulation of crack propagation using a gradient-enriched ductile damage model based on dilatational strain. *Eng. Fract. Mech.* **95**, 13–28 (2012)
- L. Malcher, F. Andrade Pires, J. Cesar de Sa, An assessment of isotropic constitutive models for ductile fracture under high and low stress triaxiality. *Int. J. Plasticity* **30–31**, 81–115 (2012)
- C. McAuliffe, H. Waisman, A unified model for metal failure capturing shear banding and fracture. *Int. J. Plasticity* **65**, 131–151 (2015)
- F. Mudry, A local approach to cleavage fracture. *Nucl. Eng. Des.* **105**, 65–76 (1987)
- S. Murakami, N. Ohno, A continuum theory of creep and creep damage, in *3rd Creep in Structures Symposium*, Leicester, pp. 422–443 (1980)
- K. Nahshon, J. Hutchinson, Modification of the gurson model to shear failure. *Eur. J. Mech. A/Solids* **27**, 1–17 (2008)
- K. Nahshon, Z. Xue, A modified gurson model and its applications to punch-out experiments. *Eng. Fract. Mech.* **76**, 997–1009 (2009)

- A. Needleman, V. Tvergaard, Dynamic crack growth in a nonlocal progressively cavitating solid. *Eur. J. Mech. A/Solids* **17**(3), 421–438 (1998)
- A. Needleman, V. Tvergaard, Numerical modeling of the ductile-brittle transition. *Int. J. Fract.* **101**, 73–97 (2000)
- T. Pardoen, J.W. Hutchinson, An extended model for void growth and coalescence, *J. Mech. Phys. Solids* **48**(12), 2467–2512 (2000)
- Y. Rabotnov, *Creep Problems in Structural Members* (North-Holland, Amsterdam, 1969)
- S. Ramaswamy, N. Aravas, Finite element implementation of gradient plasticity models. Part I: gradient-dependent yield functions. *Comput. Methods Appl. Mech. Eng.* **163**, 33–53 (1998)
- S. Renevey, S. Carassou, B. Marini, C. Eripret, A. Pineau, Ductile – brittle transition of ferritic steels modelled by the local approach to fracture. *JOURNAL DE PHYSIQUE IV Colloque C6, supplBment au Journal de Physique III* **6**, 343–352 (1996)
- R. Ritchie, J. Knott, J. Rice, On the relationship between critical tensile stress and fracture toughness in mild steel. *J. Mech. Phys. Solids* **21**, 395–410 (1973)
- F. Rivalin, A. Pineau, M. Di Fant, J. Besson, Ductile tearing of Pipeline steel wide plates-I: Dynamics and quasi static experiments. *Engng Fract. Mech* **68**(3), 329–345 (2001a)
- F. Rivalin, J. Besson, M. Di Fant, A. Pineau, Ductile tearing of Pipeline steel wide plates-II: Modeling of in-plane crack propagation. *Engng Fract. Mech* **68**(3), 347–364 (2001b)
- G. Rousselier, Ductile fracture models and their potential in local approach of fracture. *Nuc. Eng. Des.* **105**, 97–111 (1987)
- M. Samal, M. Seidenfuss, E. Roos, B.K. Dutta, H.S. Kushwaha, A mesh independent GTN damage model and its application in simulation of ductile fracture behaviour. ASME 2008 Pressure vessels and piping conference volume 3: Design and Analysis, Chicago, Illinois, USA, July 27–31 pp. 187–193 (2008)
- I. Scheider, W. Brocks, Simulation of cup-cone fracture using the cohesive model. *Eng. Fract. Mech.* **70**, 1943–1961 (2003)
- A. Shterenlikht, 3D CAFE modeling of transitional ductile-brittle fracture in steels. PhD. Thesis, Department of Mechanical Engineering, University of Sheffield (2003)
- J. Skrzypek, A. Ganczarski, F. Rustichelli, H. Egner, *Advanced Materials and Structures for Extreme Operating Conditions* (Springer, Berlin, 2008)
- C. Soyarslan, I. Turtuk, B. Deliktas, S. Bargmann, A thermomechanically consistent constitutive theory for modeling micro-void and/or micro-crack driven failure in metals at finite strains. *Int. J. Appl. Mech.* **8**, 1–20 (2016)
- S. Sutar, G. Kale, S. Merad, Analysis of ductile-to-brittle transition temperature of mild steel. *Int. J. Innov. Eng. Res. Technol.* **1**, 1–10 (2014)
- B. Tanguy, J. Besson, R. Piques, A. Pineau, Ductile to brittle transition of an a508 steel characterized by charpy impact test, Part I. Experimental results. *Eng. Fract. Mech.* **72**, 49–72 (2007)
- G. Taylor, H. Quinney, The latent energy remaining in a metal after cold working. *Proc. R. Soc. Lond.* **A143**, 307–326 (1934)
- K. Turba, B. Gulcimen, Y. Li, D. Blagoeva, P. Hähner, R. Hurst, Introduction of a new notched specimen geometry to determine fracture properties by small punch testing. *Eng. Fract. Mech.* **78**(16), 2826–2833 (2011)
- I. Turtuk, B. Deliktas, Coupled porous plasticity – continuum damage mechanics approaches in modelling temperature driven ductile-to-brittle transition fracture in ferritic steels. *Int. J. Plasticity* **77**, 246–261 (2016)
- V. Tvergaard, Influence of voids on shear band instabilities under plane strain conditions. *Int. J. Fract.* **17**, 389–407 (1981)
- V. Tvergaard, On localization in ductile materials containing spherical voids. *Int. J. Fract.* **18**, 237–252 (1982a)
- V. Tvergaard, Influence of void nucleation on ductile shear fracture at a free surface. *J. Mech. Phys. Solids* **30**, 399–425 (1982b)
- V. Tvergaard, A. Needleman, Analysis of the cup-cone fracture in a round tensile bar. *Acta Metall.* **32**, 157–169 (1984)

-
- W. Weibull, A statistical distribution function of wide applicability. *Jour. App. Mech* **18**, 293–297 (1953)
- J. Wen, Y. Huang, K. Hwang, C. Liu, M. Li, The modified gurson model accounting for the void size effect. *Int. J. Plasticity* **21**, 381–395 (2005)
- L. Xia, C. Fong Shih, A fracture model applied to the ductile/brittle regime. *Journal de Physique IV* **6**, 363–372 (1996)
- Z.L. Zhang, C. Thaulow, J. Ødegard, A complete gurson model approach for ductile fracture, *Eng. Fract. Mech* **67**(2), 155–168 (2000)



Size-Dependent Transverse Vibration of Microbeams

32

Ömer Civalek and Bekir Akgöz

Contents

Introduction	1124
Modified Strain Gradient Theory	1125
Trigonometric Shear Deformation Microbeam Model	1127
Analytical Solutions for Free Vibration Problem of Simply Supported Microbeams	1132
Numerical Results and Discussion	1133
Conclusion	1137
References	1138

Abstract

In this chapter, a new microstructure-dependent higher-order shear deformation beam model is introduced to investigate the vibrational characteristics of microbeams. This model captures both the size and shear deformation effects without the need for any shear correction factors. The governing differential equations and related boundary conditions are derived by implementing Hamilton's principle on the basis of modified strain gradient theory in conjunction with trigonometric shear deformation beam theory. The free vibration problem for simply supported microbeams is analytically solved by employing the Navier solution procedure. Moreover, a new modified shear correction factor is firstly proposed for Timoshenko (first-order shear deformation) microbeam model. Several comparative results are presented to indicate the effects of material length-scale parameter ratio, slenderness ratio, and shear correction factor on the natural frequencies of microbeams. It is observed that effect of shear deformation becomes more considerable for both smaller slenderness ratios and higher modes.

Ö. Civalek (✉) · B. Akgöz

Civil Engineering Department, Division of Mechanics, Akdeniz University, Antalya, Turkey
e-mail: civalek@yahoo.com; bekirakgoz@akdeniz.edu.tr

© Springer Nature Switzerland AG 2019

G. Z. Voyiadjis (ed.), *Handbook of Nonlocal Continuum Mechanics for Materials and Structures*, https://doi.org/10.1007/978-3-319-58729-5_8

1123

Keywords

Microbeam · Size dependency · Vibration · Small-scale effect · Modified strain gradient theory · Higher-order beam theory · Shear deformation effect · Modified shear correction factor · Length-scale parameter · Trigonometric beam model

Introduction

The miniaturized (small-sized) structures have a wide range of applications in nano- and micro-electromechanical systems (NEMS and MEMS) due to the rapid improvements in technology (Younis et al. 2003; Li and Fang 2010; Wu et al. 2010). Microbeams are one of the essential structures frequently used in MEMS/NEMS such as micro-resonators (Zook et al. 1992), atomic force microscopes (Torii et al. 1994), micro-actuators (Hung and Senturia 1999), and microswitches (Xie et al. 2003). Because of the characteristic dimensions of the microbeams (thickness, width, and length) are on the order of microns and submicrons, size effects should be taken into consideration on the determination of the mechanical characteristics of such structures. However, it has been experimentally observed for several materials that microstructural effects appear and have considerable effect on mechanical properties and deformation behavior for smaller sizes (Poole et al. 1996; Lam et al. 2003; McFarland and Colton 2005). Unfortunately, the well-known classical continuum theories, which are independent of scale of the structure's size, fail to estimate and explain of size dependency in micro- and nanoscale structures. Subsequently, various nonclassical continuum theories, which include at least one additional material length-scale parameter, have been developed like couple stress theory (Mindlin and Tiersten 1962; Koiter 1964; Toupin 1964), micropolar theory (Eringen 1967), nonlocal elasticity theory (Eringen 1972, 1983), and strain gradient theory (Fleck and Hutchinson 1993; Vardoulakis and Sulem 1995; Altan et al. 1996).

One of the higher-order continuum theories, named as strain gradient theory, developed by Fleck and Hutchinson (1993, 2001), can be viewed as extended form of the Mindlin's simplified theory (Mindlin 1965). This theory requires five additional material length-scale parameters related to second-order deformation gradients. Subsequently, Lam et al. (2003) proposed a more useful form of the strain gradient theory which is named as modified strain gradient theory (MSGT) and includes three additional material length-scale parameters for linear elastic isotropic materials.

This theory has been employed by many researchers to analyze size-dependent microbeams. For instance, Bernoulli-Euler and Timoshenko models were introduced for static bending, free vibration, and buckling behaviors of microbeams by Kong et al. (2009), Wang et al. (2010), and Akgöz and Civalek (2012, 2013a). Furthermore, Kahrobaian et al. (2012) and Ansari et al. (2011) introduced Bernoulli-Euler and Timoshenko beam models for functionally graded microbeams, respectively. Artan and Batra (2012) employed the method of initial values for the free vibration of Bernoulli-Euler strain gradient beams with four different boundary conditions as simply supported-simply supported, clamped-free, clamped-clamped,

and clamped-simply supported. Approximate solutions for static and dynamic analyses of microbeams were also carried out by finite element method based on Bernoulli-Euler and Timoshenko beam theories, respectively (Kahrobaiyan et al. 2013; Zhang et al. 2014a).

Presently, various beam theories have been proposed and used to investigate the mechanical behaviors of beams. Influences of shear deformation can be neglected for slender beams with a large aspect ratio. However, effects of shear deformation and rotary inertia become more prominent and cannot be ignored for moderately thick beams and vibration responses on higher modes. In this manner, several shear deformation beam theories have been developed to account for the effects of transverse shear. One of the earlier shear deformation beam theories is the first-order shear deformation beam theory (commonly named as Timoshenko beam theory (TBT)) (Timoshenko 1921). This theory assumes that shear stress and strain are constant along the height of the beam. In fact, the distributions of these are not uniform, and also there are no transverse shear stress and strain at the top and bottom surfaces of the beam. For this reason, a shear correction factor is needed, as a disadvantage of the theory. After that, some higher-order shear deformation beam theories, which satisfy the condition of no shear stress and strain without any shear correction factors, have been presented such as parabolic (third-order) beam theory (Levinson 1981; Reddy 1984), trigonometric (sinusoidal) beam theory (Touratier 1991), hyperbolic beam theory (Soldatos 1992), exponential beam theory (Karama et al. 2003), and general exponential beam theory (Aydogdu 2009a). These theories have been used less than Euler-Bernoulli beam theory (EBT) and TBT on prediction of the mechanical responses of microstructures on the basis of the nonclassical continuum theories (Aydogdu 2009b; Salamat-talab et al. 2012; Şimşek and Reddy 2013a, b; Thai and Vo 2012, 2013; Akgöz and Civalek 2013b, 2014a, b, c, 2015; Zhang et al. 2014b).

In the present study, a new size-dependent trigonometric (sinusoidal) shear deformation beam model in conjunction with modified strain gradient theory is developed. This model captures both the microstructural and shear deformation effects without the need for any shear correction factors. The governing differential equations and related boundary conditions are derived by using Hamilton's principle. The free vibration response of simply supported microbeams is investigated. Analytical solutions for the first three natural frequencies are presented. In order to indicate the accuracy and validity of the present model, the results are comparatively presented with the results of other beam theories. A detailed parametric study is carried out to indicate the influences of material length-scale parameter, slenderness ratio, and shear correction factors on the natural frequencies of microbeams.

Modified Strain Gradient Theory

The modified strain gradient elasticity theory was proposed by Lam et al. (2003) in which contains not only classical strain tensor but also second-order deformation gradients (first-order strain gradients) such as dilatation gradient vector and deviatoric stretch gradient and symmetric rotation gradient tensors. The strain energy U

on the basis of the modified strain gradient elasticity theory can be written by (Lam et al. 2003; Kong et al. 2009):

$$U = \frac{1}{2} \int_0^L \int_A \left(\sigma_{ij} \varepsilon_{ij} + p_i \gamma_i + \tau_{ijk}^{(1)} \eta_{ijk}^{(1)} + m_{ij}^s \chi_{ij}^s \right) dA dx \tag{1}$$

$$\varepsilon_{ij} = \frac{1}{2} (u_{i,j} + u_{j,i}) \tag{2}$$

$$\gamma_i = \varepsilon_{mm,i} \tag{3}$$

$$\eta_{ijk}^{(1)} = \frac{1}{3} (\varepsilon_{jk,i} + \varepsilon_{ki,j} + \varepsilon_{ij,k}) - \frac{1}{15} [\delta_{ij} (\varepsilon_{mm,k} + 2\varepsilon_{mk,m}) + \delta_{jk} (\varepsilon_{mm,i} + 2\varepsilon_{mi,m}) + \delta_{ki} (\varepsilon_{mm,j} + 2\varepsilon_{mj,m})] \tag{4}$$

$$\chi_{ij}^s = \frac{1}{2} (\theta_{i,j} + \theta_{j,i}) \tag{5}$$

$$\theta_i = \frac{1}{2} e_{ijk} u_{k,j} \tag{6}$$

where u_i , θ_i , ε_{ij} , γ_i , $\eta_{ijk}^{(1)}$ and χ_{ij}^s denote the components of the displacement vector \mathbf{u} , the rotation vector $\boldsymbol{\theta}$, the strain tensor $\boldsymbol{\varepsilon}$, the dilatation gradient vector $\boldsymbol{\gamma}$, the deviatoric stretch gradient tensor $\boldsymbol{\eta}^{(1)}$, and the symmetric rotation gradient tensor $\boldsymbol{\chi}^s$, respectively. Also, δ is the symbol of Kronecker delta and e_{ijk} is the permutation symbol.

Furthermore, the components of the classical stress tensor $\boldsymbol{\sigma}$ and the higher-order stress tensors \mathbf{p} , $\boldsymbol{\tau}^{(1)}$, and \mathbf{m}^s defined as (Lam et al. 2003).

$$\sigma_{ij} = \lambda \varepsilon_{mm} \delta_{ij} + 2\mu \varepsilon_{ij} \tag{7}$$

$$p_i = 2\mu l_0^2 \gamma_i \tag{8}$$

$$\tau_{ijk}^{(1)} = 2\mu l_1^2 \eta_{ijk}^{(1)} \tag{9}$$

$$m_{ij}^s = 2\mu l_2^2 \chi_{ij}^s \tag{10}$$

where l_0, l_1, l_2 are additional material length-scale parameters related to dilatation gradients, deviatoric stretch gradients, and rotation gradients, respectively. Furthermore, λ and μ are the Lamé constants defined as

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}, \quad \mu = \frac{E}{2(1+\nu)} \tag{11}$$

where E is Young's modulus and ν is Poisson's ratio.

Trigonometric Shear Deformation Microbeam Model

The displacement components of an initially straight beam on the basis of trigonometric shear deformation beam theory (see Fig. 1) can be written as (Touratier 1991).

$$\begin{aligned} u_1(x, z, t) &= u(x, t) - z \frac{\partial w(x, t)}{\partial x} + R(z)\phi(x, t) \\ u_2(x, z, t) &= 0 \\ u_3(x, z, t) &= w(x, t) \end{aligned} \quad (12)$$

in which

$$\phi(x, t) = \frac{\partial w(x, t)}{\partial x} - \varphi(x, t) \quad (13)$$

where u_1 , u_2 and u_3 are the x^- , y^- and z^- components of the displacement vector, and also u and w are the axial and transverse displacements, φ is the angle of rotation of the cross section about y^- axis of any point on the midplane of the beam, respectively. $R(z)$ is a function which depends on z and plays a role in determination of the transverse shear strain and stress distribution throughout the height of the beam. In order to satisfy no shear stress and strain condition at the upper ($z = -h/2$) and lower ($z = h/2$) surfaces of the beam, $R(z)$ is selected as following without need for any shear correction factors:

$$R(z) = \frac{h}{\pi} \sin\left(\frac{\pi z}{h}\right) \quad (14)$$

It can be noted that the displacement components for EBT and TBT will be obtained by setting $R(z)$ in Eq. 12 equal to (0) and (z) , respectively. With the use of Eqs. 12, 13, and 14 into Eq. 2, the nonzero strain components are obtained as

$$\varepsilon_{11} = \frac{\partial u}{\partial x} - z \frac{\partial^2 w}{\partial x^2} + R \frac{\partial \phi}{\partial x}, \quad \varepsilon_{13} = \frac{1}{2} S \phi \quad (15)$$

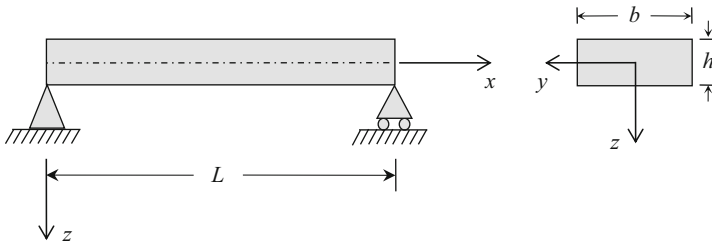


Fig. 1 Geometry, coordinate system, and cross section of a simply supported microbeam

where

$$S(z) = \cos\left(\frac{\pi z}{h}\right) \quad (16)$$

and from Eq. 15 and Eq. 3, the components of dilatation gradient vector γ are expressed as

$$\begin{aligned} \gamma_1 &= \frac{\partial^2 u}{\partial x^2} - z \frac{\partial^3 w}{\partial x^3} + R \frac{\partial^2 \phi}{\partial x^2} \\ \gamma_2 &= 0, \\ \gamma_3 &= -\frac{\partial^2 w}{\partial x^2} + S \frac{\partial \phi}{\partial x} \end{aligned} \quad (17)$$

By inserting Eq. 15 in Eq. 4, the nonzero components of deviatoric stretch gradient tensor $\eta^{(1)}$ can be obtained as

$$\begin{aligned} \eta_{111}^{(1)} &= \frac{1}{5} \left[2 \left(\frac{\partial^2 u}{\partial x^2} - z \frac{\partial^3 w}{\partial x^3} + R \frac{\partial^2 \phi}{\partial x^2} \right) + \frac{\pi^2}{h^2} R \phi \right], \\ \eta_{113}^{(1)} &= \eta_{131}^{(1)} = \eta_{311}^{(1)} = -\frac{4}{15} \left(\frac{\partial^2 w}{\partial x^2} - 2S \frac{\partial \phi}{\partial x} \right), \\ \eta_{122}^{(1)} &= \eta_{212}^{(1)} = \eta_{221}^{(1)} = -\frac{1}{5} \left(\frac{\partial^2 u}{\partial x^2} - z \frac{\partial^3 w}{\partial x^3} + R \frac{\partial^2 \phi}{\partial x^2} - \frac{\pi^2}{3h^2} R \phi \right), \\ \eta_{133}^{(1)} &= \eta_{313}^{(1)} = \eta_{331}^{(1)} = -\frac{1}{5} \left(\frac{\partial^2 u}{\partial x^2} - z \frac{\partial^3 w}{\partial x^3} + R \frac{\partial^2 \phi}{\partial x^2} + \frac{4}{3} \frac{\pi^2}{h^2} R \phi \right), \\ \eta_{223}^{(1)} &= \eta_{232}^{(1)} = \eta_{322}^{(1)} = \frac{1}{15} \left(\frac{\partial^2 w}{\partial x^2} - 2S \frac{\partial \phi}{\partial x} \right), \eta_{333}^{(1)} = \frac{1}{5} \left(\frac{\partial^2 w}{\partial x^2} - 2S \frac{\partial \phi}{\partial x} \right) \end{aligned} \quad (18)$$

Also, the use of Eq. 12 in Eq. 6 gives

$$\theta_1 = 0, \theta_2 = -\frac{\partial w}{\partial x} + \frac{1}{2} S \phi, \theta_3 = 0 \quad (19)$$

and the nonzero components of the symmetric part of the rotation gradient tensor χ^s can be achieved by using of Eq. 19 into Eq. 5 as

$$\chi_{12}^s = \chi_{21}^s = -\frac{1}{2} \left(\frac{\partial^2 w}{\partial x^2} - \frac{1}{2} S \frac{\partial \phi}{\partial x} \right), \chi_{23}^s = \chi_{32}^s = -\frac{\pi^2}{4h^2} R \phi \quad (20)$$

With the use of Eq. 15 in Eq. 7, the nonzero components of classical stress tensor σ can be written as

$$\begin{aligned} \sigma_{11} &= E \eta \left(\frac{\partial u}{\partial x} - z \frac{\partial^2 w}{\partial x^2} + R \frac{\partial \phi}{\partial x} \right), \sigma_{13} = \mu S \phi \\ \sigma_{22} &= \sigma_{33} = \frac{E\nu}{(1+\nu)(1-2\nu)} \left(\frac{\partial u}{\partial x} - z \frac{\partial^2 w}{\partial x^2} + R \frac{\partial \phi}{\partial x} \right) \end{aligned} \quad (21)$$

where

$$\eta = \frac{(1-v)}{(1+v)(1-2v)} \quad (22)$$

It is notable that Poisson's effect is neglected by choosing $\eta = 1$ in Eq. 22 (Reddy 2011). From Eq. 8 and Eq. 15, the nonzero components of higher-order stress tensor \mathbf{p} are obtained as

$$p_1 = 2\mu l_0^2 \left(\frac{\partial^2 u}{\partial x^2} - z \frac{\partial^3 w}{\partial x^3} + R \frac{\partial^2 \phi}{\partial x^2} \right), \quad p_3 = -2\mu l_0^2 \left(\frac{\partial^2 w}{\partial x^2} - S \frac{\partial \phi}{\partial x} \right) \quad (23)$$

By inserting Eq. 18 in Eq. 9, the nonzero components of higher-order stress tensor $\boldsymbol{\tau}^{(1)}$ are written as

$$\begin{aligned} \tau_{111}^{(1)} &= \frac{2}{5} \mu l_1^2 \left[2 \left(\frac{\partial^2 u}{\partial x^2} - z \frac{\partial^3 w}{\partial x^3} + R \frac{\partial^2 \phi}{\partial x^2} \right) + \frac{\pi^2}{h^2} R \phi \right], \\ \tau_{113}^{(1)} &= \tau_{131}^{(1)} = \tau_{311}^{(1)} = -\frac{8}{15} \mu l_1^2 \left(\frac{\partial^2 w}{\partial x^2} - 2S \frac{\partial \phi}{\partial x} \right), \\ \tau_{122}^{(1)} &= \tau_{212}^{(1)} = \tau_{221}^{(1)} = -\frac{2}{5} \mu l_1^2 \left(\frac{\partial^2 u}{\partial x^2} - z \frac{\partial^3 w}{\partial x^3} + R \frac{\partial^2 \phi}{\partial x^2} - \frac{\pi^2}{3h^2} R \phi \right), \\ \tau_{133}^{(1)} &= \tau_{313}^{(1)} = \tau_{331}^{(1)} = -\frac{2}{5} \mu l_1^2 \left(\frac{\partial^2 u}{\partial x^2} - z \frac{\partial^3 w}{\partial x^3} + R \frac{\partial^2 \phi}{\partial x^2} + \frac{4}{3} \frac{\pi^2}{h^2} R \phi \right), \\ \tau_{223}^{(1)} &= \tau_{232}^{(1)} = \tau_{322}^{(1)} = \frac{2}{15} \mu l_1^2 \left(\frac{\partial^2 w}{\partial x^2} - 2S \frac{\partial \phi}{\partial x} \right), \quad \tau_{333}^{(1)} = \frac{2}{5} \mu l_1^2 \left(\frac{\partial^2 w}{\partial x^2} - 2S \frac{\partial \phi}{\partial x} \right) \end{aligned} \quad (24)$$

Similarly, the nonzero components of higher-order stress tensor \mathbf{m}^s are determined by using of Eq. 20 into Eq. 10:

$$m_{12}^s = m_{21}^s = -\mu l_2^2 \left(\frac{\partial^2 w}{\partial x^2} - \frac{1}{2} S \frac{\partial \phi}{\partial x} \right), \quad m_{23}^s = m_{32}^s = -\frac{\mu l_2^2}{2} \frac{\pi^2}{h^2} R \phi \quad (25)$$

With the substitution of Eqs. 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, and 25 into Eq. 1, the first variation of strain energy of microbeam is expressed as

$$\begin{aligned} \delta U &= \int_0^L \int_A \left(\sigma_{ij} \delta \varepsilon_{ij} + p_i \delta \gamma_i + \tau_{ijk}^{(1)} \delta \eta_{ijk}^{(1)} + m_{ij}^s \delta \chi_{ij}^s \right) dA dx \\ &= \int_0^L \int_A \left(\sigma_{11} \delta \varepsilon_{11} + 2\sigma_{13} \delta \varepsilon_{13} + p_1 \delta \gamma_1 + p_3 \delta \gamma_3 + \tau_{111}^{(1)} \delta \eta_{111}^{(1)} + 3\tau_{113}^{(1)} \delta \eta_{113}^{(1)} + 3\tau_{122}^{(1)} \delta \eta_{122}^{(1)} \right. \\ &\quad \left. + 3\tau_{133}^{(1)} \delta \eta_{133}^{(1)} + 3\tau_{223}^{(1)} \delta \eta_{223}^{(1)} + \tau_{333}^{(1)} \delta \eta_{333}^{(1)} + 2m_{12}^s \delta \chi_{12}^s + 2m_{23}^s \delta \chi_{23}^s \right) dA dx \\ &= \int_0^L \left[A \left(-E \frac{\partial^2 u}{\partial x^2} + 2\mu k_1 \frac{\partial^4 u}{\partial x^4} \right) \delta u + \left(k_2 \phi - k_3 \frac{\partial^2 \phi}{\partial x^2} + k_4 \frac{\partial^3 w}{\partial x^3} + \frac{12}{\pi^2} \mu I k_1 \left(\frac{\partial^4 \phi}{\partial x^4} - \frac{4}{\pi} \frac{\partial^5 w}{\partial x^5} \right) \right) \delta \phi \right. \\ &\quad \left. + \left(k_5 \frac{\partial^4 w}{\partial x^4} - k_4 \frac{\partial^3 \phi}{\partial x^3} - 2\mu I k_1 \left(\frac{\partial^6 w}{\partial x^6} - \frac{24}{\pi^3} \frac{\partial^5 \phi}{\partial x^5} \right) \right) \delta w \right] dx + \left[A \left(E \frac{\partial u}{\partial x} - 2\mu k_1 \frac{\partial^3 u}{\partial x^3} \right) \delta u \right. \\ &\quad \left. + 2\mu A k_1 \frac{\partial^2 u}{\partial x^2} \delta \left(\frac{\partial u}{\partial x} \right) + \left((k_3 + k_6) \frac{\partial \phi}{\partial x} - (k_4 + k_7) \frac{\partial^2 w}{\partial x^2} - \frac{12}{\pi^2} \mu I k_1 \left(\frac{\partial^3 \phi}{\partial x^3} - \frac{4}{\pi} \frac{\partial^4 w}{\partial x^4} \right) \right) \delta \phi \right. \\ &\quad \left. + \left(k_6 \phi + \frac{12}{\pi^2} \mu I k_1 \left(\frac{\partial^2 \phi}{\partial x^2} - \frac{4}{\pi} \frac{\partial^3 w}{\partial x^3} \right) \right) \delta \left(\frac{\partial \phi}{\partial x} \right) + \left(k_4 \frac{\partial^2 \phi}{\partial x^2} - k_5 \frac{\partial^3 w}{\partial x^3} + 2\mu I k_1 \left(\frac{\partial^5 w}{\partial x^5} - \frac{24}{\pi^3} \frac{\partial^4 \phi}{\partial x^4} \right) \right) \delta w \right. \\ &\quad \left. + \left(k_5 \frac{\partial^2 w}{\partial x^2} - k_4 \frac{\partial \phi}{\partial x} - 2\mu I k_1 \left(\frac{\partial^4 w}{\partial x^4} - \frac{24}{\pi^3} \frac{\partial^3 \phi}{\partial x^3} \right) \right) \delta \left(\frac{\partial w}{\partial x} \right) \right. \\ &\quad \left. + \left(-k_7 \phi + 2\mu I k_1 \left(\frac{\partial^3 w}{\partial x^3} - \frac{24}{\pi^3} \frac{\partial^2 \phi}{\partial x^2} \right) \right) \delta \left(\frac{\partial^2 w}{\partial x^2} \right) \right]_0^L \end{aligned} \quad (26)$$

where L is length of the microbeam, A is the area of cross section, I is the second moment of area:

$$\begin{aligned}
 k_1 &= l_0^2 + \frac{2}{5}l_1^2, k_2 = \mu A \left(\frac{1}{2} + \frac{\pi^2}{h^2} \left(\frac{4}{15}l_1^2 + \frac{1}{8}l_2^2 \right) \right), \\
 k_3 &= \frac{6}{\pi^2}EI + \mu A \left(l_0^2 + \frac{2}{3}l_1^2 + \frac{1}{8}l_2^2 \right), \\
 k_4 &= \frac{1}{\pi} \left(\frac{24}{\pi^2}EI + \mu A \left(4l_0^2 + \frac{4}{3}l_1^2 + l_2^2 \right) \right), \\
 k_5 &= EI + \mu A \left(2l_0^2 + \frac{8}{15}l_1^2 + l_2^2 \right), \\
 k_6 &= \frac{1}{5}\mu A l_1^2, k_7 = \frac{4}{5\pi}\mu A l_1^2
 \end{aligned}
 \tag{27}$$

The kinetic energy of the microbeam is given by

$$T = \int_0^L \int_A \frac{1}{2}\rho \left[\left(\frac{\partial u_1}{\partial t} \right)^2 + \left(\frac{\partial u_2}{\partial t} \right)^2 + \left(\frac{\partial u_3}{\partial t} \right)^2 \right] dA dx
 \tag{28}$$

where ρ is the mass density. From Eqs. 12 and 28, the first variation of the kinetic energy can be expressed as

$$\begin{aligned}
 \delta T &= \int_0^L \left\{ m_0 \left[\frac{\partial u}{\partial t} \frac{\partial \delta u}{\partial t} + \frac{\partial w}{\partial t} \frac{\partial \delta w}{\partial t} \right] \right. \\
 &\quad \left. + m_2 \left[\frac{\partial^2 w}{\partial x \partial t} \frac{\partial^2 \delta w}{\partial x \partial t} - \frac{24}{\pi^3} \left(\frac{\partial \phi}{\partial t} \frac{\partial^2 \delta w}{\partial x \partial t} + \frac{\partial^2 w}{\partial x \partial t} \frac{\partial \delta \phi}{\partial t} \right) + \frac{6}{\pi^2} \frac{\partial \phi}{\partial t} \frac{\partial \delta \phi}{\partial t} \right] \right\} dx
 \end{aligned}
 \tag{29}$$

where (m_0, m_2) are the mass inertias as.

$$(m_0, m_2) = \rho \int_A (1, z^2) dA
 \tag{30}$$

The first variation of the work done by external forces can be written as

$$\begin{aligned}
 \delta W &= \int_0^L (f \delta u + q \delta w) dx + \left[\widehat{Q}_1 \delta u + \widehat{Q}_2 \delta \left(\frac{\partial u}{\partial x} \right) + \widehat{Q}_3 \delta w \right. \\
 &\quad \left. + \widehat{Q}_4 \delta \left(\frac{\partial w}{\partial x} \right) + \widehat{Q}_5 \delta \left(\frac{\partial^2 w}{\partial x^2} \right) + \widehat{Q}_6 \delta \phi + \widehat{Q}_7 \delta \left(\frac{\partial \phi}{\partial x} \right) \right]_0^L
 \end{aligned}
 \tag{31}$$

where $f(x, t)$ and $q(x, t)$ are the axial and transverse distributed loads, respectively. In addition, \widehat{Q}_i ($i = 1, 2, \dots, 7$) are the specified forces or moment of forces at the end of the microbeam. After that, with the aid of Hamilton’s principle as

$$0 = \int_0^T (\delta T - \delta U + \delta W) dt
 \tag{32}$$

and by substituting Eqs. 26, 29, and 31 into Eq. 32, integrating by parts, and setting the coefficients δu , δw , and $\delta \phi$ equal to zero, the governing equations of motion of the microbeam based on SBT can be obtained as (Akgöz and Civalek 2013b).

$$\delta u : -m_0 \frac{\partial^2 u}{\partial t^2} + A \left(E \frac{\partial^2 u}{\partial x^2} - 2\mu k_1 \frac{\partial^4 u}{\partial x^4} \right) + f = 0 \quad (33)$$

$$\delta w : -m_0 \frac{\partial^2 w}{\partial t^2} + m_2 \left(\frac{\partial^4 w}{\partial x^2 \partial t^2} - \frac{24}{\pi^3} \frac{\partial^3 \phi}{\partial x \partial t^2} \right) - k_5 \frac{\partial^4 w}{\partial x^4} \quad (34)$$

$$+ k_4 \frac{\partial^3 \phi}{\partial x^3} + 2\mu I k_1 \left(\frac{\partial^6 w}{\partial x^6} - \frac{24}{\pi^3} \frac{\partial^5 \phi}{\partial x^5} \right) + q = 0$$

$$\delta \phi : \frac{24}{\pi^3} m_2 \frac{\partial^3 w}{\partial x \partial t^2} - \frac{6}{\pi^2} m_2 \frac{\partial^2 \phi}{\partial t^2} - k_2 \phi + k_3 \frac{\partial^2 \phi}{\partial x^2} \quad (35)$$

$$- k_4 \frac{\partial^3 w}{\partial x^3} - \frac{12}{\pi^2} \mu I k_1 \left(\frac{\partial^4 \phi}{\partial x^4} - \frac{4}{\pi} \frac{\partial^5 w}{\partial x^5} \right) = 0$$

and boundary conditions at $x = 0$ and $x = L$

$$\text{either } A \left(E \frac{\partial u}{\partial x} - 2\mu k_1 \frac{\partial^3 u}{\partial x^3} \right) = \widehat{Q}_1 \text{ or } u = 0 \quad (36)$$

$$\text{either } 2\mu A k_1 \frac{\partial^2 u}{\partial x^2} = \widehat{Q}_2 \text{ or } \frac{\partial u}{\partial x} = 0 \quad (37)$$

$$\text{either } -k_5 \frac{\partial^3 w}{\partial x^3} + k_4 \frac{\partial^2 \phi}{\partial x^2} + 2\mu I k_1 \left(\frac{\partial^5 w}{\partial x^5} - \frac{24}{\pi^3} \frac{\partial^4 \phi}{\partial x^4} \right) \quad (38)$$

$$- m_2 \left(\frac{24}{\pi^3} \frac{\partial^2 \phi}{\partial t^2} - \frac{\partial^3 w}{\partial x \partial t^2} \right) = \widehat{Q}_3 \text{ or } w = 0$$

$$\text{either } k_5 \frac{\partial^2 w}{\partial x^2} - k_4 \frac{\partial \phi}{\partial x} - 2\mu I k_1 \left(\frac{\partial^4 w}{\partial x^4} - \frac{24}{\pi^3} \frac{\partial^3 \phi}{\partial x^3} \right) = \widehat{Q}_4 \text{ or } \frac{\partial w}{\partial x} = 0 \quad (39)$$

$$\text{either } -k_7 \phi + 2\mu I k_1 \left(\frac{\partial^3 w}{\partial x^3} - \frac{24}{\pi^3} \frac{\partial^2 \phi}{\partial x^2} \right) = \widehat{Q}_5 \text{ or } \frac{\partial^2 w}{\partial x^2} = 0 \quad (40)$$

$$\text{either } (k_3 + k_6) \frac{\partial \phi}{\partial x} - (k_4 + k_7) \frac{\partial^2 w}{\partial x^2} - \frac{12}{\pi^2} \mu I k_1 \left(\frac{\partial^3 \phi}{\partial x^3} - \frac{4}{\pi} \frac{\partial^4 w}{\partial x^4} \right) = \widehat{Q}_6 \text{ or } \phi = 0 \quad (41)$$

$$\text{either } k_6 \phi + \frac{12}{\pi^2} \mu I k_1 \left(\frac{\partial^2 \phi}{\partial x^2} - \frac{4}{\pi} \frac{\partial^3 w}{\partial x^3} \right) = \widehat{Q}_7 \text{ or } \frac{\partial \phi}{\partial x} = 0 \quad (42)$$

Analytical Solutions for Free Vibration Problem of Simply Supported Microbeams

Here, in order to solve free vibration problem of simply supported microbeams, the Navier solution procedure is used. The well-known geometric boundary conditions for a simply supported end can be defined as zero deflection and nonzero slope and/or rotation of the cross section as

$$w = 0, \quad \frac{\partial w}{\partial x} \neq 0, \quad \phi \neq 0 \quad (43)$$

In view of Eq. 43, the left sides of Eqs. 39 and 41 must vanish. Hence, the following relations can be written by Eqs. 36, 37, 38, 39, 40, 41, 42, and 43 as

$$\frac{\partial u}{\partial x} = 0, \quad \frac{\partial^2 w}{\partial x^2} = 0, \quad \frac{\partial \phi}{\partial x} = 0, \quad \widehat{Q}_1 = 0, \quad \widehat{Q}_4 = 0, \quad \widehat{Q}_6 = 0 \quad (44)$$

The following expansions of generalized displacements which include undetermined Fourier coefficients and certain trigonometric functions can be successfully employed as

$$w(x, t) = \sum_{n=1}^{\infty} W_n \sin \alpha x e^{i\omega_n t} \quad (45)$$

$$\phi(x, t) = \sum_{n=1}^{\infty} H_n \cos \alpha x e^{i\omega_n t} \quad (46)$$

where W_n and H_n are the undetermined Fourier coefficients, ω_n is natural frequency, and $\alpha = \frac{n\pi}{L}$. This means that Eqs. 45 and 46 must satisfy the corresponding boundary conditions. Substituting Eqs. 45 and 46 into Eqs. 35 and 36 as the governing equations for free vibration, the following equation is obtained as

$$\left(\begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} - \omega^2 \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \right) \begin{Bmatrix} H_n \\ W_n \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \quad (47)$$

where.

$$\begin{aligned} K_{11} &= k_2 + \alpha^2 k_3 + \alpha^4 \frac{12}{\pi^2} \mu I k_1, & K_{12} &= K_{21} = -\alpha^3 \left(k_4 + \alpha^2 \frac{48}{\pi^3} \mu I k_1 \right), \\ K_{22} &= \alpha^4 (k_5 + \alpha^2 2\mu I k_1) \\ M_{11} &= \frac{6}{\pi^2} m_2, & M_{12} &= M_{21} = -\alpha \frac{24}{\pi^3} m_2, & M_{22} &= (m_0 + \alpha^2 m_2) \end{aligned} \quad (48)$$

For a nontrivial solution, the determinant of coefficient matrix must be vanished and the characteristic equation can be reached by providing this condition. The eigenvalues are obtained by solving the characteristic equation. It can be noted that the smallest root of the characteristic equation gives the first natural (fundamental) frequency.

Numerical Results and Discussion

In this section, free vibration problem of a simply supported microbeam is analytically solved with the Navier-type solution based on trigonometric shear deformation beam theory in conjunction with modified strain gradient theory. For illustration purpose, the microbeam is taken to be made of epoxy with the following material properties: Young's modulus $E = 1.44$ GPa, Poisson's ratio $\nu = 0.38$, the mass density $\rho = 1,220$ kg/m³ and the material length-scale parameter $l = 11.01$ μ m (Kahrobaian et al. 2013). The microbeam has a rectangular cross section, and the width-to-thickness ratio is taken to be constant as $b/h = 2$, while the length-to-thickness ratio is taken several values as $L/h = 5 \sim 80$. All material length-scale parameters are considered to be equal to each other as $l_0 = l_1 = l_2 = l$.

As stated before, Timoshenko beam theory (TBT) needs a shear correction factor to take into consideration the nonuniformity of transverse shear strain and stress throughout the beam thickness. For rectangular cross-section beams, the most commonly used shear correction factors can be defined as $k_s = 5/6$ (used here) and $k_s = (5 + 5\nu)/(6 + 5\nu)$. The classical results evaluated by TBT and other shear deformation beam theories such as third-order (parabolic), trigonometric (sinusoidal), hyperbolic, and exponential shear deformation beam theories are in good agreement. However, this agreement may decrease for the results of higher-order continuum theories, and this situation can be seen from the previous works (Akgöz and Civalek 2013b; Şimşek and Reddy 2013a, b). Consequently, a new modified shear correction factor (k_s^*) is used for Timoshenko microbeam model (TBT*)-based MSGT as follows (Akgöz and Civalek 2014a):

$$k_s^* = k_s k_{ac}^{MSGT} \quad (49)$$

where

$$k_{ac}^{MSGT} = 15 \left(\frac{l_0 + l_1 + l_2}{3} \right)^a / h^a \quad a = 3 \left(\frac{h}{\left(\frac{l_0 + l_1 + l_2}{3} \right)} \right)^{0.08} - 0.45 \quad (50)$$

It can be noted that k_s^* will be equal to k_s by setting material length-scale parameters equal to zero in Eq. 50. In order to demonstrate the accuracy and validity of the present analysis, some illustrative examples are comparatively given with other beam theories.

Table 1 Dimensionless fundamental frequencies ($\bar{\omega}_1 = \omega_1 L^2 \sqrt{m_0/EI}$)

l/h	Beam theory	$L = 8 h$		$L = 40 h$		$L = 80 h$	
		CT	MSGT	CT	MSGT	CT	MSGT
0	EBT	9.8696	9.8696	9.8696	9.8696	9.8696	9.8696
	TBT	9.6094	9.6094	9.8587	9.8587	9.8669	9.8669
	TBT*	9.6094	9.6094	9.8587	9.8587	9.8669	9.8669
	SBT	9.6098	9.6098	9.8587	9.8587	9.8669	9.8669
0.5	EBT	9.8696	21.8020	9.8696	21.7179	9.8696	21.7153
	TBT	9.6094	19.7861	9.8587	21.6223	9.8669	21.6913
	TBT*	9.6094	21.2380	9.8587	21.6944	9.8669	21.7094
	SBT	9.6098	21.2186	9.8587	21.6933	9.8669	21.7091
1	EBT	9.8696	40.1133	9.8696	39.9305	9.8696	39.9248
	TBT	9.6094	31.3909	9.8587	39.3696	9.8669	39.7817
	TBT*	9.6094	38.9701	9.8587	39.8826	9.8669	39.9128
	SBT	9.6098	39.0201	9.8587	39.8843	9.8669	39.9132

*Timoshenko beam model with the new shear correction factor

Table 2 Dimensionless second natural frequencies ($\bar{\omega}_2 = \omega_2 L^2 \sqrt{m_0/EI}$)

l/h	Beam theory	$L = 8 h$		$L = 40 h$		$L = 80 h$	
		CT	MSGT	CT	MSGT	CT	MSGT
0	EBT	39.4784	39.4784	39.4784	39.4784	39.4784	39.4784
	TBT	35.8237	35.8237	39.3048	39.3048	39.4348	39.4348
	TBT*	35.8237	35.8237	39.3048	39.3048	39.4348	39.4348
	SBT	35.8329	35.8329	39.3050	39.3050	39.4348	39.4348
0.5	EBT	39.4784	88.2502	39.4784	86.9138	39.4784	86.8718
	TBT	35.8237	66.0744	39.3048	85.4193	39.4348	86.4892
	TBT*	35.8237	80.2054	39.3048	86.5394	39.4348	86.7776
	SBT	35.8329	80.1416	39.3050	86.5229	39.4348	86.7733
1	EBT	39.4784	162.7167	39.4784	159.8137	39.4784	159.7222
	TBT	35.8237	95.8638	39.3048	151.4933	39.4348	157.4783
	TBT*	35.8237	146.5687	39.3048	159.052	39.4348	159.5306
	SBT	35.8329	147.5984	39.3050	159.0795	39.4348	159.5373

*Timoshenko beam model with the new shear correction factor

Dimensionless first three natural frequencies for various values of l/h and slenderness ratios corresponding to different beam theories are tabulated in Tables 1, 2, and 3, respectively. It can be clearly observed from the tables that the dimensionless natural frequencies predicted by both CT and TBT are lower than the other ones, while those obtained by both MSGT and EBT are larger than the other ones. Also, an increase in l/h leads to an increment in the difference between dimensionless natural frequencies corresponding to classical and nonclassical models, and also this difference becomes more prominent for higher modes. On the other hand,

Table 3 Dimensionless third natural frequencies ($\bar{\omega}_3 = \omega_3 L^2 \sqrt{m_0/EI}$)

l/h	Beam theory	$L = 8 h$		$L = 40 h$		$L = 80 h$	
		CT	MSGT	CT	MSGT	CT	MSGT
0	EBT	88.8264	88.8264	88.8264	88.8264	88.8264	88.8264
	TBT	73.2989	73.2989	87.9565	87.9565	88.6060	88.6060
	TBT*	73.2989	73.2989	87.9565	87.9565	88.6060	88.6060
	SBT	73.3581	73.3581	87.9576	87.9576	88.6062	88.6062
0.5	EBT	88.8264	202.4110	88.8264	195.7139	88.8264	195.5009
	TBT	73.2989	126.2815	87.9565	188.4270	88.6060	193.5834
	TBT*	73.2989	167.4196	87.9565	193.8353	88.6060	195.0255
	SBT	73.3581	168.1073	87.9576	193.7568	88.6062	195.0041
1	EBT	88.8264	374.4469	88.8264	359.9240	88.8264	359.4607
	TBT	73.2989	186.4415	87.9565	322.3023	88.6060	348.4635
	TBT*	73.2989	304.9236	87.9565	356.1056	88.6060	358.4933
	SBT	73.3581	310.7247	87.9576	356.2506	88.6062	358.5275

*Timoshenko beam model with the new shear correction factor

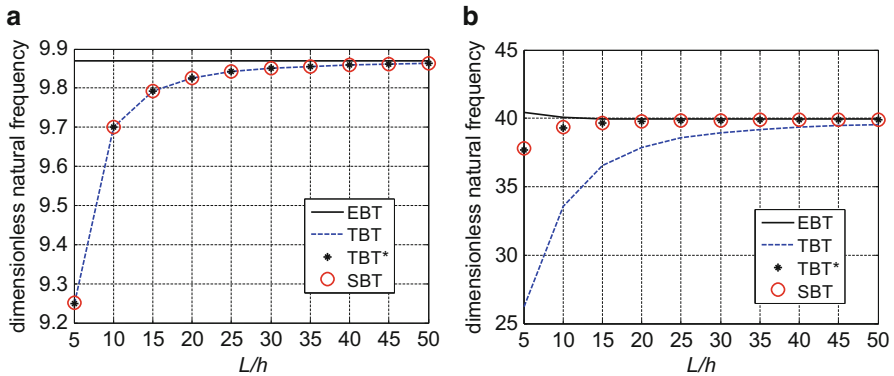


Fig. 2 Variations of the dimensionless natural frequency versus slenderness ratio (first mode). (a) CT (b) MSGT

difference between the results corresponding to EBT and shear deformation beam theories (TBT, TBT*, and SBT) is more significant for short beams. This situation can be interpreted as the effect of shear deformation is minor for slender beams with a large slenderness ratio. In addition, it can be clearly seen from the tables that the natural frequencies predicted by SBT and TBT* are in good agreement, while the divergence between the natural frequencies of SBT and TBT is considerable especially for bigger values of l/h .

Variations of the dimensionless first three natural frequencies of the simply supported microbeam with respect to the slenderness ratio corresponding to different beam models are depicted in Figs. 2, 3, and 4, respectively. It is observed that an

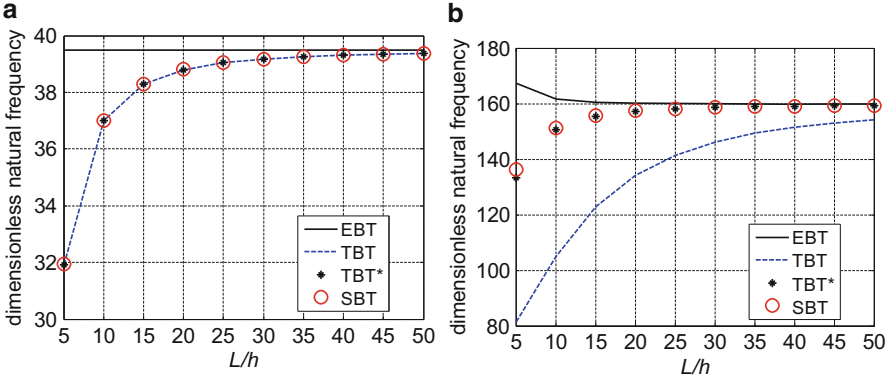


Fig. 3 Variations of the dimensionless natural frequency versus slenderness ratio (second mode). (a) CT (b) MSGT

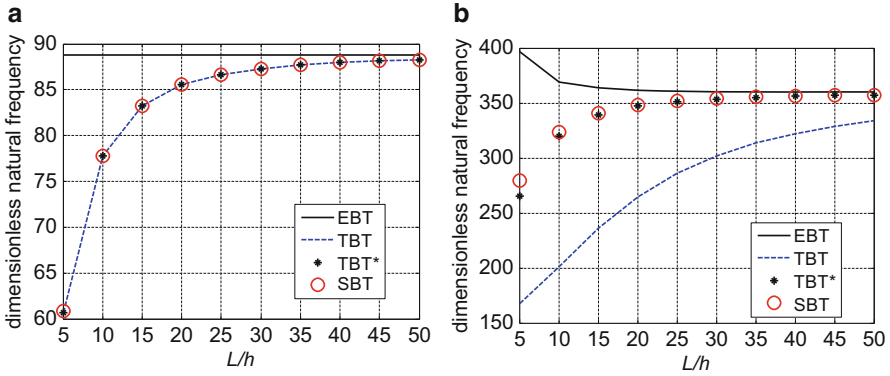


Fig. 4 Variations of the dimensionless natural frequency versus slenderness ratio (third mode). (a) CT (b) MSGT

increase in slenderness ratio leads to a decrement on effects of shear deformation, and differences between the dimensionless natural frequencies based on EBT, TBT, TBT*, and SBT are diminishing for $L/h \geq 50$. Moreover, it can be concluded that the dimensionless natural frequencies evaluated by TBT, TBT*, and SBT are nearly equal to each other for CT, but the difference between TBT and SBT is more considerable in the higher-order models for lower slenderness ratios and higher modes.

Influences of h/l ratio on the first three dimensionless natural frequencies for $L = 7h$ are illustrated in Figs. 5, 6, and 7, respectively. These figures reveal that natural frequencies based on MSGT are always bigger than CT. Also, it is found that the effects of shear deformation and small size are more considerable for smaller values of h/l and higher modes.

Fig. 5 Effects of thickness-to-material length-scale parameter ratio on the first dimensionless natural frequency ($L = 7h$)

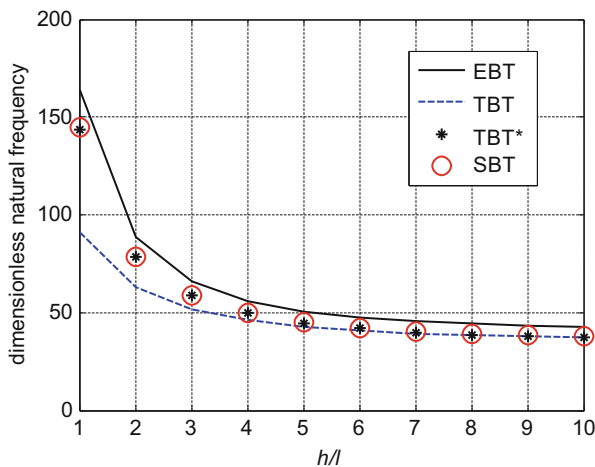
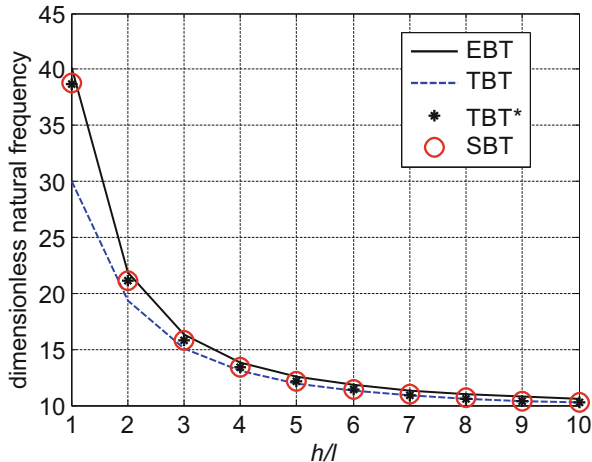
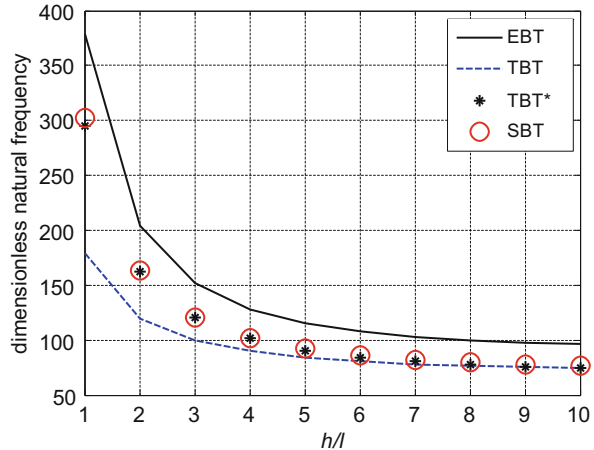


Fig. 6 Effects of thickness-to-material length-scale parameter ratio on the second dimensionless natural frequency ($L = 7h$)

Conclusion

In this study, a size-dependent sinusoidal shear deformation beam model in conjunction with modified strain gradient elasticity theory (MSGT) is developed. The model captures both the microstructural and shear deformation effects without any shear correction factors. The governing differential equations and corresponding boundary conditions are derived by using Hamilton’s principle. The free vibration behavior of simply supported microbeams is investigated. Analytical solutions for the first three natural frequencies are presented by the Navier solution technique. The results are compared with other beam theories for the validation of the

Fig. 7 Effects of thickness-to-material length-scale parameter ratio on the third dimensionless natural frequency ($L = 7h$)



present model. A detailed parametric study is carried out to show the influences of thickness-to-material length-scale parameter ratio, slenderness ratio, and shear deformation on the free vibration response of simply supported microbeams. The obtained results can be summarized as:

- Microbeams based on MSGT are stiffer than based on the classical theory.
- The natural frequencies obtained by both MSGT and EBT are always greater than those predicted by the other considered beam models and theories.
- The difference between the natural frequencies decreases as the thickness-to-material length-scale parameter ratio increases.
- Effect of shear deformation becomes more considerable for both smaller slenderness ratios and higher modes.
- Use of modified shear correction factors is more suitable for Timoshenko microbeam models based on higher-order continuum theories.

Acknowledgments This study has been supported by The Scientific and Technological Research Council of Turkey (TÜBİTAK) with Project No: 112M879. This support is gratefully acknowledged.

References

- B. Akgöz, Ö. Civalek, Arch. Appl. Mech. **82**, 423 (2012)
 B. Akgöz, Ö. Civalek, Acta Mech. **224**, 2185 (2013a)
 B. Akgöz, Ö. Civalek, Int. J. Eng. Sci. **70**, 1 (2013b)
 B. Akgöz, Ö. Civalek, Compos. Struct. **112**, 214 (2014a)
 B. Akgöz, Ö. Civalek, Int. J. Mech. Sci. **81**, 88 (2014b)
 B. Akgöz, Ö. Civalek, Int. J. Eng. Sci. **85**, 90 (2014c)
 B. Akgöz, Ö. Civalek, Int. J. Mech. Sci. **99**, 10 (2015)
 B.S. Altan, H. Evensen, E.C. Aifantis, Mech. Res. Commun. **23**, 35 (1996)

- R. Ansari, R. Gholami, S. Sahmani, *Compos. Struct.* **94**, 221 (2011)
- R. Artan, R.C. Batra, *Acta Mech.* **223**, 2393 (2012)
- M. Aydogdu, *Compos. Struct.* **89**, 94 (2009a)
- M. Aydogdu, *Phys. E.* **41**, 1651 (2009b)
- A.C. Eringen, *Z. Angew. Math. Phys.* **18**, 12 (1967)
- A.C. Eringen, *Int. J. Eng. Sci.* **10**, 1 (1972)
- A.C. Eringen, *J. Appl. Phys.* **54**, 4703 (1983)
- N.A. Fleck, J.W. Hutchinson, *J. Mech. Phys. Solids* **41**, 1825 (1993)
- N.A. Fleck, J.W. Hutchinson, *J. Mech. Phys. Solids* **49**, 2245 (2001)
- E.S. Hung, S.D. Senturia, *J. Microelectromech. Syst.* **8**, 497 (1999)
- M.H. Kahrobaian, M. Asghari, M.T. Ahmadian, *Finite Elem. Anal. Des.* **68**, 63 (2013)
- M.H. Kahrobaian, M. Rahaeifard, S.A. Tajalli, M.T. Ahmadian, *Int. J. Eng. Sci.* **52**, 65 (2012)
- M. Karama, K.S. Afaq, S. Mistou, *Int. J. Solids Struct.* **40**, 1525 (2003)
- W.T. Koiter, *Proc. K. Ned. Akad. Wet. B* **67**, 17 (1964)
- S. Kong, S. Zhou, Z. Nie, K. Wang, *Int. J. Eng. Sci.* **47**, 487 (2009)
- D.C.C. Lam, F. Yang, A.C.M. Chong, J. Wang, P. Tong, *J. Mech. Phys. Solids* **51**, 1477 (2003)
- M. Levinson, A new rectangular beam theory. *J. Sound Vib.* **74**, 81 (1981)
- P. Li, Y. Fang, *J. Micromech. Microeng.* **20**, 035005 (2010)
- A.W. McFarland, J.S. Colton, *J. Micromech. Microeng.* **15**, 1060 (2005)
- R.D. Mindlin, *Int. J. Solids Struct.* **1**, 417 (1965)
- R.D. Mindlin, H.F. Tiersten, *Arch. Ration. Mech. Anal.* **11**, 415 (1962)
- W.J. Poole, M.F. Ashby, N.A. Fleck, *Scr. Mater.* **34**, 559 (1996)
- J.N. Reddy, *J. Appl. Mech.* **51**, 745 (1984)
- J.N. Reddy, *J. Mech. Phys. Solids* **59**, 2382 (2011)
- M. Salamat-talab, A. Nateghi, J. Torabi, *Int. J. Mech. Sci.* **57**, 63 (2012)
- M. Şimşek, J.N. Reddy, *Int. J. Eng. Sci.* **64**, 37 (2013a)
- M. Şimşek, J.N. Reddy, *Compos. Struct.* **101**, 47 (2013b)
- K.P. Soldatos, *Acta Mech.* **94**, 195 (1992)
- H.T. Thai, T.P. Vo, *Int. J. Eng. Sci.* **54**, 58 (2012)
- H.T. Thai, T.P. Vo, *Compos. Struct.* **96**, 376 (2013)
- S.P. Timoshenko, *Philos. Mag.* **41**, 744 (1921)
- A. Torii, M. Sasaki, K. Hane, S. Okuma, *Sensors Actuators A Phys.* **44**, 153 (1994)
- R.A. Toupin, *Arch. Ration. Mech. Anal.* **17**, 85 (1964)
- M. Touratier, *Int. J. Eng. Sci.* **29**, 901 (1991)
- I. Vardoulakis, J. Sulem, *Bifurcation Analysis in Geomechanics* (Blackie/Chapman and Hall, London, 1995)
- B. Wang, J. Zhao, S. Zhou, *Eur. J. Mech. A/Solids* **29**, 591 (2010)
- Z.Y. Wu, H. Yang, X.X. Li, Y.L. Wang, *J. Micromech. Microeng.* **20**, 115014 (2010)
- W.C. Xie, H.P. Lee, S.P. Lim, *Nonlinear Dyn.* **31**, 243 (2003)
- M.I. Younis, E.M. Abdel-Rahman, A.H. Nayfeh, *J. Microelectromech. Syst.* **12**, 672 (2003)
- B. Zhang, Y. He, D. Liu, Z. Gan, L. Shen, *Finite Elem. Anal. Des.* **79**, 22 (2014a)
- B. Zhang, Y. He, D. Liu, Z. Gan, L. Shen, *Eur J Mech A/Solids* **47**, 211 (2014b)
- J.D. Zook, D.W. Burns, H. Guckel, J.J. Sniogowski, R.L. Engelstad, Z. Feng, *Sensors Actuators A Phys.* **35**, 51 (1992)



Axial Vibration of Strain Gradient Micro-rods

33

Ömer Civalek, Bekir Akgöz, and Babür Deliktaş

Contents

Introduction	1142
Formulation for Modified Strain Gradient Theory	1144
Microstructure-Dependent Rod Model	1145
Solution of Axial Vibration Problem	1148
Numerical Results and Discussion	1152
Conclusion	1153
References	1154

Abstract

In this chapter, size-dependent axial vibration response of micro-sized rods is investigated on the basis of modified strain gradient elasticity theory. On the contrary to the classical rod model, the developed nonclassical micro-rod model includes additional material length scale parameters and can capture the size effect. If the additional material length scale parameters are equal to zero, the current model reduces to the classical one. The equation of motion together with initial conditions, classical and nonclassical corresponding boundary conditions, for micro-rods is derived by implementing Hamilton's principle. The resulting higher-order equation is analytically solved for clamped-free and clamped-clamped boundary conditions. Finally, some illustrative examples are presented

Ö. Civalek (✉) · B. Akgöz

Civil Engineering Department, Division of Mechanics, Akdeniz University, Antalya, Turkey

e-mail: civalek@yahoo.com; bekirakgoz@akdeniz.edu.tr

B. Deliktaş

Faculty of Engineering-Architecture, Department of Civil Engineering, Uludag University, Bursa, Turkey

e-mail: bdeliktas@uludag.edu.tr

to indicate the influences of the additional material length scale parameters, size dependency, boundary conditions, and mode numbers on the natural frequencies. It is found that size effect is more significant when the micro-rod diameter is closer to the additional material length scale parameter. In addition, it is observed that the difference between natural frequencies evaluated by the present and classical models becomes more considerable for both lower values of slenderness ratio and higher modes.

Keywords

Micro-rod · Size dependency · Axial vibration · Small-scale effect · Modified strain gradient theory · Length scale parameter · Higher-order rod model · Natural frequency

Introduction

Nowadays, due to the rapid advances in technologies, micro- and nano-sized mechanical systems like microbeams, microbars, biosensors, nanowires, atomic force microscope, nanotubes, micro actuators, nano probes, micro- and nano-electromechanical systems (MEMS and NEMS), and ultra-thin films have been widely used in modern applications such as mechanical, biomedical, chemical, and biological applications (Fu et al. 2003; Li et al. 2003; Najar et al. 2005; Faris and Nayfeh 2007; Moser and Gijs 2007; Kahrobaiyan et al. 2011a). The insight of the mechanical behavior characteristics of micro- and nanostructures is very important for the optimum design of such structures. Bending, buckling, and vibration responses of these structures can be investigated by experimental studies and computer simulation techniques at atomistic levels. The effects of size dependency on the deformation behaviors of the aforementioned structures have been experimentally observed (Fleck et al. 1994; Chong and Lam 1999; Senturia 2001; Haque and Saif 2003; Lam et al. 2003; Lou et al. 2006).

Due to the difficulty and computationally expensiveness of experimentation and simulation techniques at atomistic levels (e.g., molecular dynamic simulation), many scientists and researchers tended the continuum mechanics modeling as an alternative. However, the classical continuum mechanics approaches do not the ability for interpretation and explanation of the microstructural dependency of small-sized structures due to the lack of any additional (intrinsic) material length scale parameters. Then, higher-order (nonclassical) continuum theories, which include at least one additional material length scale parameter in addition to classical ones, have been proposed to predict the microstructure-dependent behavior of these small-scale structures.

Higher-order continuum theories include Cosserat elasticity by Cosserat and Cosserat (1909), strain gradient elasticity of Mindlin (1964,1965), micropolar theory (Eringen and Suhubi 1964), nonlocal elasticity (Eringen 1983), couple stress theory by Mindlin and Tiersten (1962), Toupin (1962), and Koiter (1964), strain

gradient theory (Fleck and Hutchinson 1993), simple gradient elasticity with surface energy (Vardoulakis and Sulem 1995; Altan et al. 1996; Altan and Aifantis 1997), modified couple stress (Yang et al. 2002), and modified strain gradient theories (Lam et al. 2003). Some earlier studies based on these theories available in the literature have been briefly given here.

Peddiesson et al. (2003) formulated a nonlocal Bernoulli-Euler beam model with nonlocal elasticity theory. Also, Reddy (2007a) investigated bending, buckling, and free vibration analysis of nonlocal beams for different beam theories. Wang et al. (2008) studied bending problem of micro- and nano-sized beams based on nonlocal Timoshenko beam theory. Aydogdu (2009) investigated the small-scale effect on longitudinal vibration of a nanorod on the basis of Eringen's nonlocal elasticity theory.

The classical couple stress theory has been used to investigate the bending analysis of a circular cylinder by Anthoine (2000). Tsepoura et al. (2002) investigated static and dynamic analysis of bars based on simple gradient elasticity theory with surface energy. Papargyri-Beskou et al. (2003a) and Lazopoulos (2012) observed dynamic analysis of gradient elastic beams. Bending and buckling analysis of gradient elastic beams is studied on the basis of Bernoulli-Euler beam model by Papargyri-Beskou et al. (2003b) and Lazopoulos and Lazopoulos (2010).

Modified couple stress theory is a higher-order continuum theory that has been elaborated by Yang et al. (2002) which contains the symmetric rotation gradient tensor and one additional material length scale parameter in addition to the conventional (classical) strain tensor. Park and Gao (2006) and Ma et al. (2008) developed new size-dependent Bernoulli-Euler and Timoshenko beam models, respectively. Kong et al. (2008) investigated free vibration analysis of the Bernoulli-Euler microbeam model based on this theory.

The modified strain gradient elasticity theory is one of the popular higher-order continuum theories, which was proposed by Lam et al. (2003), that includes dilatation and deviatoric stretch gradient tensors besides the symmetric rotation gradient and classical strain tensor and also three additional material length scale parameters for linear elastic isotropic materials. Kong et al. (2009), Akgöz and Civalek (2011), and Wang et al. (2010) used modified strain gradient elasticity for static and dynamic analyses of microbeams on the basis of Bernoulli-Euler and Timoshenko beam models, respectively. Furthermore, static torsion and torsional free vibration analyses of microbars based on this theory were presented by Kahrobaiyan et al. (2011b) and Narendar et al. (2012). Recently, longitudinal vibration responses of microbars were investigated by Akgöz and Civalek (2013, 2014), Kahrobaiyan et al. (2013), and Güven (2014).

In this chapter, size-dependent axial vibration response of micro-sized rods is investigated on the basis of modified strain gradient elasticity theory. The equation of motion together with initial conditions, classical and nonclassical corresponding boundary conditions, for micro-rods is derived with the aid of Hamilton's principle. The resulting higher-order equation is solved for two different boundary conditions as clamped-free and clamped-clamped. Influences of micro-rod characteristic

lengths, slenderness ratio, additional material length scale parameters, and mode number on the vibrational response of the size-dependent micro-rod are investigated.

Formulation for Modified Strain Gradient Theory

The strain energy U in a linear elastic isotropic material occupying volume V based on the modified strain gradient elasticity theory can be written by Lam et al. (2003) and Kong et al. (2009).

$$U = \frac{1}{2} \int_V \left(\sigma_{ij} \varepsilon_{ij} + p_i \gamma_i + \tau_{ijk}^{(1)} \eta_{ijk}^{(1)} + m_{ij}^s \chi_{ij}^s \right) dv \quad (1)$$

$$\varepsilon_{ij} = \frac{1}{2} (u_{i,j} + u_{j,i}) \quad (2)$$

$$\gamma_i = \varepsilon_{mm,i} \quad (3)$$

$$\eta_{ijk}^{(1)} = \frac{1}{3} (\varepsilon_{jk,i} + \varepsilon_{ki,j} + \varepsilon_{ij,k}) - \frac{1}{15} [\delta_{ij} (\varepsilon_{mm,k} + 2\varepsilon_{mk,m}) + \delta_{jk} (\varepsilon_{mm,i} + 2\varepsilon_{mi,m}) + \delta_{ki} (\varepsilon_{mm,j} + 2\varepsilon_{mj,m})] \quad (4)$$

$$\chi_{ij}^s = \frac{1}{2} (\theta_{i,j} + \theta_{j,i}) \quad (5)$$

$$\theta_i = \frac{1}{2} e_{ijk} u_{k,j} \quad (6)$$

where u_i , ε_{ij} , γ_i , $\eta_{ijk}^{(1)}$, χ_{ij}^s , and θ_i are the components of the displacement vector, the strain tensor, the dilatation gradient vector, the deviatoric stretch gradient tensor, the symmetric rotation gradient tensor, and the rotation vector, respectively. Also δ_{ij} is the Kronecker delta, and e_{ijk} is the permutation symbol. The stress measures σ_{ij} , p_i , $\tau_{ijk}^{(1)}$, and m_{ij}^s are the components of classical and higher-order stresses defined as Lam et al. (2003).

$$\sigma_{ij} = \lambda \delta_{ij} \varepsilon_{mm} + 2\mu \varepsilon_{ij} \quad (7)$$

$$p_i = 2\mu l_0^2 \gamma_i \quad (8)$$

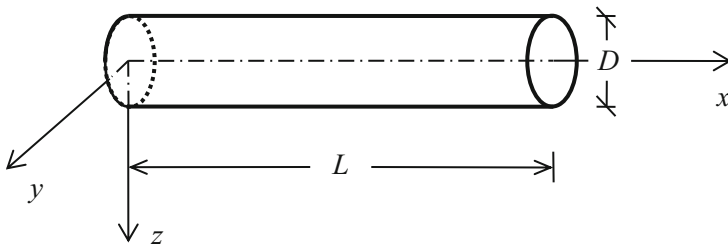


Fig. 1 Geometry and coordinate system of a straight micro-rod

$$\tau_{ijk}^{(1)} = 2\mu l_1^2 \eta_{ijk}^{(1)} \tag{9}$$

$$m_{ij}^s = 2\mu l_2^2 \chi_{ij}^s \tag{10}$$

where λ and μ are the well-known Lamé constants and l_0, l_1, l_2 are additional material length scale parameters which represent the size dependency and related to dilatation gradients, deviatoric stretch gradients, and rotation gradients, respectively.

Microstructure-Dependent Rod Model

It is considered that the case of axial vibration of a straight thin micro-rod (see Fig. 1). Due to axial vibrations take place in x - direction, the deformation of the cross section in y - and z - directions is assumed to be negligible by a simple theory for axial vibration of thin rods. The components of displacement vector can be expressed as Rao (2007).

$$u_1 = u(x, t), u_2 = 0, u_3 = 0 \tag{11}$$

where $u_1, u_2,$ and u_3 are the components of displacement vector in x -, y -, and z - directions, respectively.

In view of Eqs. (2) and (11), the non-zero strain component of the micro-rod is

$$\varepsilon_{xx} = \frac{\partial u}{\partial x} \tag{12}$$

Use of Eqs. (12) into (3), the non-zero component of dilatation gradient vector is obtained as

$$\gamma_x = \frac{\partial^2 u}{\partial x^2} \tag{13}$$

From Eqs. (2) and (4), non-zero components of deviatoric stretch gradient tensor are achieved as

$$\eta_{xxx}^{(1)} = \frac{2}{5} \frac{\partial^2 u}{\partial x^2}, \eta_{xyy}^{(1)} = \eta_{xzz}^{(1)} = \eta_{yxy}^{(1)} = \eta_{yyx}^{(1)} = \eta_{zxx}^{(1)} = \eta_{zzx}^{(1)} = -\frac{1}{5} \frac{\partial^2 u}{\partial x^2} \quad (14)$$

Furthermore, all components of rotation vector and so symmetric rotation gradient tensor are equal to zero as

$$\theta_i = 0, (i = x, y, z) \quad (15)$$

$$\chi_{ij}^s = 0, (i, j = x, y, z) \quad (16)$$

The non-zero stress σ_{ij} can be obtained by neglecting Poisson's effect in Eq. (7) as

$$\sigma_{xx} = E \frac{\partial u}{\partial x} \quad (17)$$

where E is the elastic modulus. By inserting Eqs. (13) into (8) and Eqs. (14) into (9), the non-zero components of higher-order stresses p_i and $\tau_{ijk}^{(1)}$ can be expressed as

$$p_x = 2\mu l_0^2 \frac{\partial^2 u}{\partial x^2} \quad (18)$$

$$\tau_{xxx}^{(1)} = \frac{4}{5} \mu l_1^2 \frac{\partial^2 u}{\partial x^2}, \tau_{xyy}^{(1)} = \tau_{xzz}^{(1)} = \tau_{yxy}^{(1)} = \tau_{yyx}^{(1)} = \tau_{zxx}^{(1)} = \tau_{zzx}^{(1)} = -\frac{2}{5} \mu l_1^2 \frac{\partial^2 u}{\partial x^2}, \quad (19)$$

Substituting above equations into Eq. (1), the strain energy U can be rewritten as

$$U = \frac{1}{2} \int_0^L \left\{ EA(u')^2 + \left(2\mu Al_0^2 + \frac{4}{5} \mu Al_1^2 \right) (u'')^2 \right\} dx \quad (20)$$

where A is the cross-sectional area of the micro-rod and

$$u' = \frac{\partial u}{\partial x}, u'' = \frac{\partial^2 u}{\partial x^2} \quad (21)$$

The first variation of strain energy U in Eq. (20) on the time interval $[t_0, t_1]$ can be calculated as following expression

$$\delta U = \int_{t_0}^{t_1} \int_0^L \{Bu^{(4)} - EAu''\} \delta u dx dt + \int_{t_0}^{t_1} \left\{ [(EAu' - Bu''') \delta u]_0^L + [Bu'' \delta u']_0^L \right\} dt \tag{22}$$

where

$$u''' = \frac{\partial^3 u}{\partial x^3}, u^{(4)} = \frac{\partial^4 u}{\partial x^4}, B = 2\mu A \left(l_0^2 + \frac{2}{5} l_1^2 \right) \tag{23}$$

On the other hand, the first variation of the work done by external force q , axial force N , and higher-order axial force N^h on the time interval $[t_0, t_1]$ takes the following form

$$\delta W = \int_{t_0}^{t_1} \int_0^L q \delta u dx dt + \int_{t_0}^{t_1} \left\{ [N \delta u]_0^L + [N^h \delta u']_0^L \right\} dt \tag{24}$$

Also, the first variation of kinetic energy K of the micro-rod on the time interval $[t_0, t_1]$ reads as

$$\delta K = - \int_{t_0}^{t_1} \int_0^L m \ddot{u} \delta u dx dt + \int_0^L [m \dot{u} \delta u]_{t_0}^{t_1} dx \tag{25}$$

where m is the mass per unit length and

$$\dot{u} = \frac{\partial u}{\partial t}, \ddot{u} = \frac{\partial^2 u}{\partial t^2} \tag{26}$$

The following relation is written by employing Hamilton’s principle with Eqs. (22, 24, and 25)

$$\delta \left\{ \int_{t_0}^{t_1} [K - (U - W)] dt \right\} = 0 \tag{27}$$

$$\int_{t_0}^{t_1} \int_0^L [EAu'' - Bu^{(4)} - m\ddot{u} + q] \delta u dx dt + \int_{t_0}^{t_1} \left\{ [N - (EAu' - Bu''')] \delta u \right\}_0^L dt + \int_{t_0}^{t_1} \left\{ [N^h - Bu''] \delta u' \right\}_0^L dt + \int_0^L [m \dot{u} \delta u]_{t_0}^{t_1} dx = 0 \tag{28}$$

According to the fundamental lemma of calculus of variation (Reddy 2007b), the equation of motion for the micro-rod reads as

$$EAu'' - Bu^{(4)} + q = m\ddot{u} \tag{29}$$

Also, initial conditions and boundary conditions satisfy the following equations, respectively, as

$$\dot{u}(x, t_1) \delta u(x, t_1) - \dot{u}(x, t_0) \delta u(x, t_0) = 0 \quad (30)$$

$$\begin{aligned} & \left[N(L, t) - \left(EAu'(L, t) - Bu'''(L, t) \right) \right] \delta u(L, t) \\ & - \left[N(0, t) - \left(EAu'(0, t) - Bu'''(0, t) \right) \right] \delta u(0, t) = 0 \end{aligned} \quad (31)$$

$$\left[N^h(L, t) - Bu''(L, t) \right] \delta u'(L, t) - \left[N^h(0, t) - Bu''(0, t) \right] \delta u'(0, t) = 0 \quad (32)$$

Solution of Axial Vibration Problem

u can be expressed as the following form by employing separation of variables method

$$u(x, t) = U(x)e^{i\omega t} \quad (33)$$

By substituting above equation into Eq. (29) in the absence of q yields

$$B \frac{d^4 U}{dx^4} - EA \frac{d^2 U}{dx^2} - \omega^2 m U = 0 \quad (34)$$

Analytical solution of Eq. (34) can be obtained as follows

$$U(x) = D_1 \sin \alpha x + D_2 \cos \alpha x + D_3 \sinh \beta x + D_4 \cosh \beta x \quad (35)$$

where

$$\alpha = \left(\frac{-EA + \sqrt{(EA)^2 + 4Bm\omega^2}}{2B} \right)^{1/2}, \beta = \left(\frac{EA + \sqrt{(EA)^2 + 4Bm\omega^2}}{2B} \right)^{1/2} \quad (36)$$

and D_i ($i = 1, 2, 3, 4$) are constants which can be determined by corresponding boundary conditions. For a micro-rod that both ends are clamped, classical and nonclassical boundary conditions are

$$U(0) = 0 \text{ and } U(L) = 0 \quad (37)$$

$$BU''(0) = 0 \text{ and } BU''(L) = 0 \quad (38)$$

By using of above boundary conditions in Eq. (35), solution can be written in a matrix form as

$$\begin{bmatrix} 0 & 1 & 0 & 1 \\ \sin \alpha L & \cos \alpha L & \sinh \beta L & \cosh \beta L \\ 0 & -\alpha^2 & 0 & \beta^2 \\ -\alpha^2 \sin \alpha L & -\alpha^2 \cos \alpha L & \beta^2 \sinh \beta L & \beta^2 \cosh \beta L \end{bmatrix} \begin{bmatrix} D_1 \\ D_2 \\ D_3 \\ D_4 \end{bmatrix} = 0 \quad (39)$$

For a nontrivial solution, the determinant of coefficient matrix of above equation must be vanished. This leads to the following condition as

$$\sin \alpha L = 0 \text{ namely } \alpha = \frac{n\pi}{L}, (n = 1, 2, \dots) \quad (40)$$

By inserting Eqs. (40) in (36), the natural longitudinal frequencies of a clamped-clamped micro-rod are obtained as

$$\omega_n = \frac{n\pi}{L} \sqrt{\frac{1}{m} \left(EA + B \frac{n^2 \pi^2}{L^2} \right)} \quad (41)$$

For a clamped-free micro-rod, classical and nonclassical boundary conditions are

$$U(0) = 0 \text{ and } EAU'(L) - BU'''(L) = 0 \quad (42)$$

$$BU''(0) = 0 \text{ and } U'(L) = 0 \quad (43)$$

By using of above boundary conditions in Eq. (35), solution can be given in a matrix form as

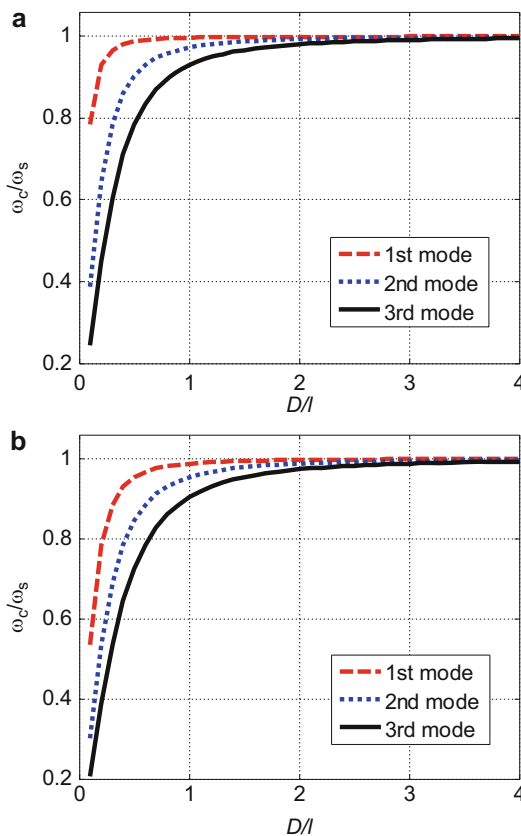
Table 1 Comparison of dimensionless natural frequencies ($\bar{\omega} = \omega L \sqrt{\frac{\rho}{E}}$) of clamped-free micro-rod for the first three modes with various values of l/D

l/D	Mode 1		Mode 2		Mode 3	
	CT	MSGT	CT	MSGT	CT	MSGT
0	1.5708	1.5708	4.7124	4.7124	7.8540	7.8540
0.3	1.5708	1.5712	4.7124	4.7243	7.8540	7.9091
0.6	1.5708	1.5726	4.7124	4.7599	7.8540	8.0721
0.9	1.5708	1.5748	4.7124	4.8187	7.8540	8.3368
1.2	1.5708	1.5779	4.7124	4.8998	7.8540	8.6938
1.5	1.5708	1.5818	4.7124	5.0021	7.8540	9.1323

Table 2 Comparison of dimensionless natural frequencies ($\omega = \omega L \sqrt{\frac{\rho}{E}}$) of clamped-clamped micro-rod for the first three modes with various values of l/D

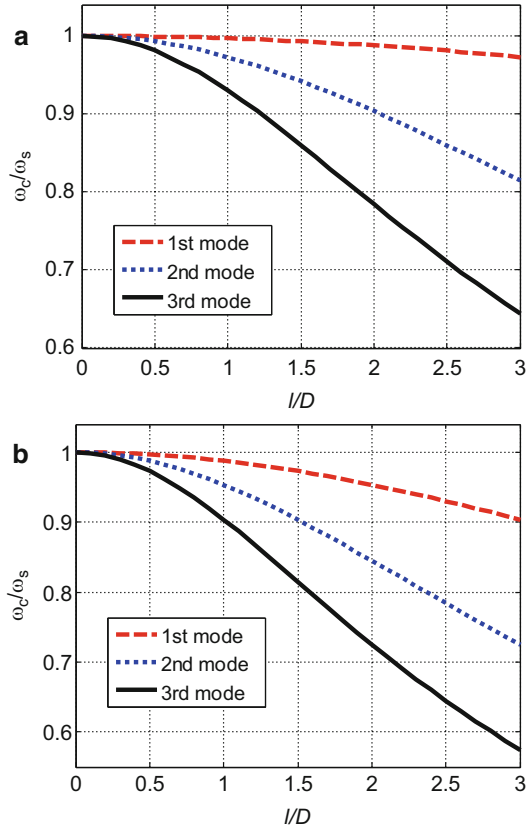
l/D	Mode 1		Mode 2		Mode 3	
	CT	MSGT	CT	MSGT	CT	MSGT
0	3.1416	3.1416	6.2832	6.2832	9.4248	9.4248
0.3	3.1416	3.1451	6.2832	6.3114	9.4248	9.5198
0.6	3.1416	3.1557	6.2832	6.3954	9.4248	9.7995
0.9	3.1416	3.1733	6.2832	6.5330	9.4248	10.2487
1.2	3.1416	3.1977	6.2832	6.7209	9.4248	10.8463
1.5	3.1416	3.2288	6.2832	6.9550	9.4248	11.5694

Fig. 2 Variations of the frequency ratio with respect to D/l for the first three modes. (a) Clamped-free (b) clamped-clamped



$$\begin{bmatrix} 0 & 1 & 0 & 1 \\ \alpha \lambda_1 \cos \alpha L & -\alpha \lambda_1 \sin \alpha L & \beta \lambda_2 \cosh \beta L & \beta \lambda_2 \sinh \beta L \\ 0 & -\alpha^2 & 0 & \beta^2 \\ \alpha \cos \alpha L & -\alpha \sin \alpha L & \beta \cosh \beta L & \beta \sinh \beta L \end{bmatrix} \begin{bmatrix} D_1 \\ D_2 \\ D_3 \\ D_4 \end{bmatrix} = 0 \quad (44)$$

Fig. 3 Variations of the first three frequency ratios versus l/D . (a) Clamped-free (b) clamped-clamped



where

$$\lambda_1 = (EA + \alpha^2 B), \lambda_2 = (EA - \beta^2 B) \tag{45}$$

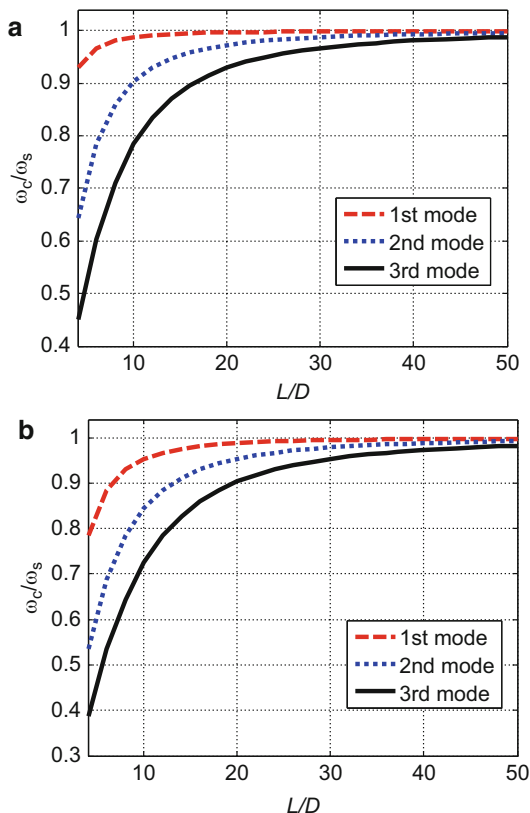
Similarly, the determinant of coefficient matrix of Eq. (44) must be vanished for a nontrivial solution. This leads to the following condition as

$$\cos \alpha L = 0 \text{ namely } \alpha = \frac{(2n - 1) \pi}{2L}, (n = 1, 2, \dots) \tag{46}$$

By inserting Eqs. (46) in (36), the natural longitudinal frequencies of a clamped-free micro-rod are achieved as

$$\omega_n = \frac{(2n - 1) \pi}{2L} \sqrt{\frac{1}{m} \left(EA + B \frac{(2n - 1)^2 \pi^2}{4L^2} \right)} \tag{47}$$

Fig. 4 Effect of slenderness ratio of the micro-rod on the frequency ratios for the first three modes ($l = D$). (a) Clamped-free (b) clamped-clamped



It is evident that if the additional material length scale parameters l_0 and l_1 are equal to zero, the natural longitudinal frequencies ω_n in Eqs. (41) and (47) will be transformed those in classical theory.

Numerical Results and Discussion

In this section, some illustrative examples for clamped-free and clamped-clamped micro-sized rods are presented. In the figures, the natural longitudinal frequencies obtained by MSGT and CT represented by ω_s and ω_c , respectively, and unless otherwise stated, $L = 20D$, $l_0 = l_1 = l$ are considered, and Poisson’s ratio is chosen as 0.38.

A comparison of nondimensional first three natural frequencies of the micro-rod corresponding to various values of the additional material length scale parameters-to-diameter ratio is tabulated in Tables 1 and 2 for clamped-free and clamped-clamped boundary conditions, respectively. It is seen from the tables that the results of classical and newly developed model are identical for $l/D = 0$. It can be said

that the nondimensional natural frequencies obtained by the new model increase gradually for bigger values of l/D , while those obtained by a classical model are not affected by the variation in l/D . It is also notable that differences between the results of the classical model and the current model are more prominent for larger values of l/D and higher modes.

Variations of the frequency ratios (ω_c / ω_s) with respect to D/l for the first three modes are illustrated in the Fig. 2 for clamped-free and clamped-clamped boundary conditions, respectively. It is evident that an increase in the values of D/l leads to an increment in the frequency ratios, and the frequency ratios are nearly equal to one for $D/l \geq 4$. Also, higher values of ω_c / ω_s are obtained for lower modes. It can be concluded that the divergence between classical and size-dependent frequencies becomes more significant for higher modes.

Variations of the first three frequency ratios are plotted versus l/D in Fig. 3 for clamped-free and clamped-clamped micro-rods, respectively. When the values of l/D increases, the frequency ratios for first, second, and third modes decrease. Also, it is noted that the frequency ratios are equal one for $l/D = 0$.

Influences of slenderness ratio on the frequency ratios are depicted for the first three modes in Fig. 4. It can be interpreted that the difference between natural frequencies predicted by the newly developed and classical models becomes more prominent for both lower slenderness ratios and higher modes. In addition, it can be said that the size dependency of the micro-rod diminishes due to an increase in the slenderness ratio.

It can be seen clearly from the present results that additional material length scale parameters are more important both clamped-free and clamped-clamped cases for smaller sizes and higher modes. Also, the values of the frequency ratios for clamped-clamped boundary condition are smaller than those of the other case.

Conclusion

A higher-order continuum theory is used for modeling of longitudinal vibration problem of micro-sized rod. Some parametric results have been presented in order to show the effect of additional length scale parameters. The results of modified strain gradient theory (MSGT) compared with those obtained by classical theory (CT). It has been seen that the frequency ratios decrease when l/D increases. It is also shown that the length scale parameters have some notable influences on axial vibration of the micro-sized rod. It is also possible to say that the effect of the length scale parameters is more significant for slender rods. An increase in slenderness ratio of the micro-rod leads to a decrease in the difference between natural frequencies predicting by the newly developed and classical models. It is also observed that additional material length scale parameters play an important role for smaller size of the micro-rod and higher modes. It is also notable that when additional material length scale parameters are zero, the present model directly becomes the classical model.

Acknowledgments This study has been supported by The Scientific and Technological Research Council of Turkey (TÜBİTAK) with Project No: 112 M879. This support is gratefully acknowledged.

References

- B. Akgöz, Ö. Civalek, *Int. J. Eng. Sci.* **49**, 1268 (2011)
- B. Akgöz, Ö. Civalek, *Compos. Part B* **55**, 263 (2013)
- B. Akgöz, Ö. Civalek, *J. Vib. Control.* **20**, 606 (2014)
- B.S. Altan, E.C. Aifantis, *J. Mech. Behav. Mater.* **8**, 231 (1997)
- B.S. Altan, H. Evensen, E.C. Aifantis, *Mech. Res. Commun.* **23**, 35 (1996)
- A. Anthoine, *Int. J. Solids Struct.* **37**, 1003 (2000)
- M. Aydogdu, *Phys. E.* **41**, 861 (2009)
- A.C.M. Chong, D.C.C. Lam, *J. Mater. Res.* **14**, 4103 (1999)
- E. Cosserat, F. Cosserat, *Theory of Deformable Bodies* (Trans. by D.H. Delphenich) (Scientific Library, A. Hermann and Sons, Paris, 1909)
- A.C. Eringen, *J. Appl. Phys.* **54**, 4703 (1983)
- A.C. Eringen, E.S. Suhubi, *Int. J. Eng. Sci.* **2**, 189 (1964)
- W. Faris, A.H. Nayfeh, *Commun. Nonlinear Sci. Numer. Simul.* **12**, 776 (2007)
- N.A. Fleck, J.W. Hutchinson, *J. Mech. Phys. Solids* **41**, 1825 (1993)
- N.A. Fleck, G.M. Muller, M.F. Ashby, J.W. Hutchinson, *Acta Metall. Mater.* **42**, 475 (1994)
- Y.Q. Fu, H.J. Du, S. Zhang, *Mater. Lett.* **57**, 2995 (2003)
- U. Güven, *Comptes Rendus Mécanique* **342**, 8 (2014)
- M.A. Haque, M.T.A. Saif, *Acta Mater.* **51**, 3053 (2003)
- M.H. Kahrobaian, M. Asghari, M.T. Ahmadian, *Int. J. Eng. Sci.* **44**, 66–67 (2013)
- M.H. Kahrobaian, M. Rahaeifard, M.T. Ahmadian, *Appl. Math. Model.* **35**, 5903 (2011a)
- M.H. Kahrobaian, S.A. Tajalli, M.R. Movahhedy, J. Akbari, M.T. Ahmadian, *Int. J. Eng. Sci.* **49**, 856 (2011b)
- W.T. Koiter, *Proc. K. Ned. Akad. Wet. B* **67**, 17 (1964)
- S. Kong, S. Zhou, Z. Nie, K. Wang, *Int. J. Eng. Sci.* **46**, 427 (2008)
- S. Kong, S. Zhou, Z. Nie, K. Wang, *Int. J. Eng. Sci.* **47**, 487 (2009)
- D.C.C. Lam, F. Yang, A.C.M. Chong, J. Wang, P. Tong, *J. Mech. Phys. Solids* **51**, 1477 (2003)
- A.K. Lazopoulos, *Int. J. Mech. Sci.* **58**, 27 (2012)
- K.A. Lazopoulos, A.K. Lazopoulos, *Eur. J. Mech. A/Solids* **29**, 837 (2010)
- X. Li, B. Bhushan, K. Takashima, C.W. Baek, Y.K. Kim, *Ultramicroscopy* **97**, 481 (2003)
- J. Lou, P. Shrotriya, S. Allameh, T. Buchheit, W.O. Soboyejo, *Mater. Sci. Eng. A* **441**, 299 (2006)
- H.M. Ma, X.L. Gao, J.N. Reddy, *J. Mech. Phys. Solids* **56**, 3379 (2008)
- R.D. Mindlin, *Arch. Ration. Mech. Anal.* **16**, 51 (1964)
- R.D. Mindlin, *Int. J. Solids Struct.* **1**, 417 (1965)
- R.D. Mindlin, H.F. Tiersten, *Arch. Ration. Mech. Anal.* **11**, 415 (1962)
- Y. Moser, M.A.M. Gijs, *Miniaturized flexible temperature sensor. J. Microelectromech. Syst.* **16**, 1349 (2007)
- F. Najjar, S. Choura, S. El-Borgi, E.M. Abdel-Rahman, A.H. Nayfeh, *J. Micromech. Microeng.* **15**, 419 (2005)
- S. Narendar, S. Ravinder, S. Gopalakrishnan, *Int. J. Nano Dimens.* **3**, 1 (2012)
- S. Papargyri-Beskou, D. Polyzos, D.E. Beskos, *Struct. Eng. Mech.* **15**, 705 (2003a)
- S. Papargyri-Beskou, K.G. Tsepoura, D. Polyzos, D.E. Beskos, *Int. J. Solids Struct.* **40**, 385 (2003b)
- S.K. Park, X.L. Gao, *J. Micromech. Microeng.* **16**, 2355 (2006)
- J. Peddieson, G.R. Buchanan, R.P. McNitt, *Int. J. Eng. Sci.* **41**, 305 (2003)
- S.S. Rao, *Vibration of Continuous Systems* (Wiley Inc, Hoboken, 2007)
- J.N. Reddy, *Int. J. Eng. Sci.* **45**, 288–307 (2007a)

- J.N. Reddy, *Theory and Analysis of Elastic Plates and Shells*, 2nd edn. (Taylor & Francis, Philadelphia, 2007b)
- S.D. Senturia, *Microsystem Design* (Kluwer Academic Publishers, Boston, 2001)
- R.A. Toupin, *Arch. Ration. Mech. Anal.* **11**, 385 (1962)
- K.G. Tsepoura, S. Papargyri-Beskou, D. Polyzos, D.E. Beskos, *Arch. Appl. Mech.* **72**, 483 (2002)
- I. Vardoulakis, J. Sulem, *Bifurcation Analysis in Geomechanics*, Blackie (Chapman and Hall, London, 1995)
- C.M. Wang, S. Kitipornchai, C.W. Lim, M. Eisenberger, *J. Eng. Mech.* **134**, 475 (2008)
- B. Wang, J. Zhao, S. Zhou, *Eur. J. Mech. A/Solids* **29**, 591 (2010)
- F. Yang, A.C.M. Chong, D.C.C. Lam, P. Tong, *Int. J. Solids Struct.* **39**, 2731 (2002)

Part V
Peridynamics



S. A. Silling

Contents

Purpose of the Peridynamic Theory	1160
Basic Concepts	1161
Properties of States	1163
Balance of Momentum	1164
Energy Balance and Thermodynamics	1166
Material Models	1167
Thermodynamic Form of a Material Model	1169
Restriction on the Heat Transport Model	1171
Damage as a Thermodynamic Variable	1173
Examples of Material Models	1174
Bond-Based Linear Material with Damage	1174
Ordinary State-Based Linear Material with Damage	1175
Non-ordinary State-Based Material	1176
Bond-Based Viscoelastic Material	1176
Isotropic Bond-Based Material	1176
Nonconvex Bond-Based Material	1177
Discrete Particles as Peridynamic Materials	1177
Effectively Eulerian Material Models	1179
Plasticity	1180
Damage Evolution	1183
Bond Breakage	1187
Connections with the Local Theory	1188
Local Kinematics and Kinetics	1188
Convergence of Peridynamics to the Local Theory	1190
Local Continuum Damage Mechanics	1192
Discussion	1193
References	1194

S. A. Silling (✉)
Sandia National Laboratories, Albuquerque, NM, USA
e-mail: sasilli@sandia.gov

Abstract

The peridynamic theory is a nonlocal extension of continuum mechanics that is compatible with the physical nature of cracks as discontinuities. It avoids the need to evaluate the partial derivatives of the deformation with respect to the spatial coordinates, instead using an integro-differential equation for the linear momentum balance. This chapter summarizes the peridynamic theory, emphasizing the continuum mechanical and thermodynamic aspects. Formulation of material models is discussed, including details on the statement of models using mathematical objects called *peridynamic states* that are nonlocal and nonlinear generalizations of second-order tensors. Damage evolution is treated within a nonlocal thermodynamic framework making use of the dependence of free energy on damage. Continuous, stable growth of damage can suddenly become unstable, leading to dynamic fracture. Peridynamics treats fracture and long-range forces on the same mathematical basis as continuous deformation and contact forces, extending the applicability of continuum mechanics to new classes of problems.

Keywords

Peridynamic · Nonlocal · Damage · Elasticity · Plasticity · Eulerian

Purpose of the Peridynamic Theory

In spite of its many successes, the local theory of continuum mechanics has some limitations that have hindered its applicability to many important problems:

- Its equations cannot be applied directly on a growing discontinuity in the deformation, making it impossible to model fracture using these equations alone.
- It does not include long-range forces such as electrostatic and van der Waals forces that are important in many technologies.
- It cannot be applied to the mechanics of discrete particles, creating a fundamental divide between molecular dynamics and continuum mechanics.

The peridynamic theory addresses these limitations in the local theory. Its field equations are integro-differential equations that do not require a smooth deformation, allowing fracture to be modeled on the same basis as continuous deformation. It treats all internal forces as long range, allowing interactions such as electrostatic forces to be included in a material description in a natural way. Discrete particles can be treated as a type of peridynamic material, allowing continuous media and systems of particles to be included within the same model, following the same basic equations.

The general idea is that the peridynamic theory replaces the local equilibrium equation with a nonlocal expression as follows:

$$\nabla \cdot \boldsymbol{\sigma} + \mathbf{b} = 0 \quad \rightarrow \quad \int_{\mathcal{H}_x} \mathbf{f}(\mathbf{q}, \mathbf{x}) dV_{\mathbf{q}} + \mathbf{b} = 0, \quad (1)$$

where $\boldsymbol{\sigma}$ is the stress field, \mathbf{b} is the body force density field, and \mathbf{f} is a vector field representing the force density (per unit volume squared) that material point \mathbf{q} exerts on \mathbf{x} . \mathcal{H}_x is a neighborhood of \mathbf{x} to be described below.

The main purposes of this chapter are as follows:

- To explain the origin of the second of (Eq. 1)
- To show how the deformation determines \mathbf{f} through the material model
- To describe material models within the framework of nonlocal thermodynamics
- To demonstrate that damage and fracture fit naturally into this framework
- To describe how some aspects of the local theory can be obtained as a limiting case of the peridynamic equations

Basic Concepts

In a peridynamic body \mathcal{B} , each material point \mathbf{x} interacts directly with its neighbors $\mathbf{q} \in \mathcal{B}$ located within a cutoff distance δ of \mathbf{x} in the reference configuration. Let \mathbf{q} denote such a neighbor and define the *bond* $\boldsymbol{\xi} = \mathbf{q} - \mathbf{x}$. The *family* of \mathbf{x} , denoted by \mathcal{H}_x , consists of all the bonds with length no greater than δ :

$$\mathcal{H}_x = \{\mathbf{q} - \mathbf{x} : 0 < |\mathbf{q} - \mathbf{x}| \leq \delta, \mathbf{q} \in \mathcal{B}\}.$$

The cutoff distance δ is called the *horizon* and is assumed for purposes of this discussion to be independent of \mathbf{x} (Fig. 1).

Let $\mathbf{y}(\mathbf{x})$ denote the deformation. Suppose there is a strain-energy density field $W(\mathbf{x})$, and that its value at \mathbf{x} depends on the *collective deformation* of \mathcal{H}_x . This means that $W(\mathbf{x})$ depends not on $\partial\mathbf{y}/\partial\mathbf{x}$ (which may not exist if a crack is present) but on $\mathbf{y}(\mathbf{q})$ for all the material points within the horizon of \mathbf{x} .

To express the dependence of W on the collective deformation of \mathcal{H}_x , it is convenient to use mathematical objects called *states*. A state is simply a mapping from \mathcal{H}_x to some other quantity, which can be a scalar, a vector, or a tensor. By convention, the bond that the state operates on is written in angle brackets, $\underline{\mathbf{A}} \langle \boldsymbol{\xi} \rangle$. State-valued fields depend on position and possibly time, denoted by

$$\underline{\mathbf{A}}[\mathbf{x}, t] \langle \boldsymbol{\xi} \rangle.$$

The fundamental kinematical quantity for purposes of material modeling is the *deformation state* $\underline{\mathbf{Y}}$, defined by:

$$\underline{\mathbf{Y}}[\mathbf{x}] \langle \mathbf{q} - \mathbf{x} \rangle = \mathbf{y}(\mathbf{q}) - \mathbf{y}(\mathbf{x}) \quad \forall \mathbf{q} \in \mathcal{H}_x, \quad \forall \mathbf{x} \in \mathcal{B}$$

Geometrically, the deformation state maps each bond to its image under the deformation.

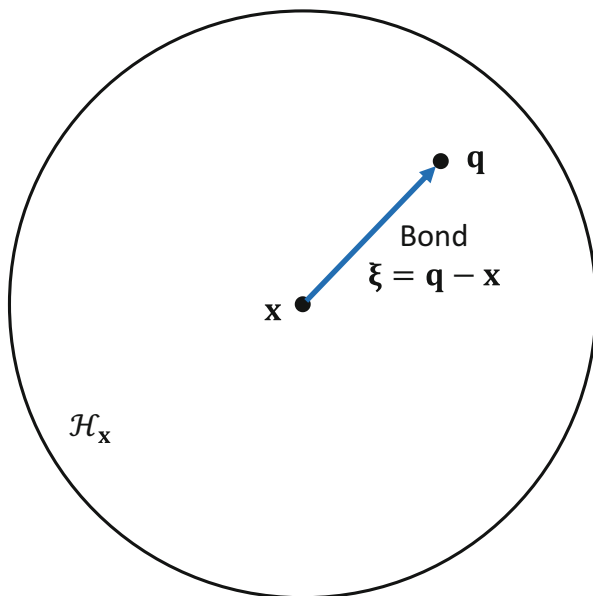


Fig. 1 Bond and family

Returning to our strain-energy density field $W(\mathbf{x})$, we can now represent its dependence on the collective deformation of \mathcal{H}_x . We write this function in the form:

$$W(\mathbf{x}) = \widehat{W}(\underline{\mathbf{Y}}[x]) \quad \forall x \in \mathcal{B}$$

where $\widehat{W}(\underline{\mathbf{Y}})$ is the *strain-energy density function* for the material. The dependence of \widehat{W} on $\underline{\mathbf{Y}}$ contains all the material-dependent characteristics of the model. Examples of such a material model are given by:

- Isotropic *bond-based* material:

$$\widehat{W}(\underline{\mathbf{Y}}) = \frac{1}{2} \int_{\mathcal{H}_x} C(|\xi|) (|\underline{\mathbf{Y}}(\xi)| - |\xi|)^2 dV_\xi,$$

where C is a scalar-valued function of bond length.

- A possible model for a fluid (one of many, see section “[Effectively Eulerian Material Models](#)”):

$$\widehat{W}(\underline{\mathbf{Y}}) = \frac{1}{2} \left(\int_{\mathcal{H}_x} C(|\xi|) (|\underline{\mathbf{Y}}(\xi)| - |\xi|) dV_\xi \right)^2.$$

Properties of States

Vector-valued states are similar to second-order tensors in that they both map vectors onto vectors. However, states can be nonlinear and even discontinuous mappings. It is useful to define the *dot product* of two vector-valued states $\underline{\mathbf{A}}$ and $\underline{\mathbf{B}}$ by:

$$\underline{\mathbf{A}} \bullet \underline{\mathbf{B}} = \int_{\mathcal{H}_x} \underline{\mathbf{A}} \langle \boldsymbol{\xi} \rangle \cdot \underline{\mathbf{B}} \langle \boldsymbol{\xi} \rangle dV_{\boldsymbol{\xi}} \quad (2)$$

and of two scalar-valued states \underline{a} and \underline{b} by:

$$\underline{a} \bullet \underline{b} = \int_{\mathcal{H}_x} \underline{a} \langle \boldsymbol{\xi} \rangle \underline{b} \langle \boldsymbol{\xi} \rangle dV_{\boldsymbol{\xi}}.$$

The *unit state* $\underline{\mathbf{1}}$ and *zero state* $\underline{\mathbf{0}}$ are defined by:

$$\underline{\mathbf{1}} \langle \boldsymbol{\xi} \rangle = 1, \quad \underline{\mathbf{0}} \langle \boldsymbol{\xi} \rangle = 0 \quad \forall \boldsymbol{\xi}.$$

The *identity state* leaves bonds unchanged:

$$\underline{\mathbf{X}} \langle \boldsymbol{\xi} \rangle = \boldsymbol{\xi} \quad \forall \boldsymbol{\xi}.$$

The *norm* of a state $\underline{\mathbf{A}}$ is defined by:

$$\|\underline{\mathbf{A}}\| = \sqrt{\underline{\mathbf{A}} \bullet \underline{\mathbf{A}}} \quad \text{or} \quad \|\underline{a}\| = \sqrt{\underline{a} \bullet \underline{a}}.$$

The composition of two states $\underline{\mathbf{A}}$ and $\underline{\mathbf{B}}$ is a state defined by:

$$(\underline{\mathbf{A}} \circ \underline{\mathbf{B}}) \langle \boldsymbol{\xi} \rangle = \underline{\mathbf{A}} \langle \underline{\mathbf{B}} \langle \boldsymbol{\xi} \rangle \rangle;$$

The “ \circ ” symbol distinguishes this operation from the point product, to be defined below. If \mathbf{R} is a second-order tensor, it may similarly be combined with a state $\underline{\mathbf{A}}$; in this case, a special symbol is not necessary:

$$(\underline{\mathbf{A}}\mathbf{R}) \langle \boldsymbol{\xi} \rangle = \underline{\mathbf{A}} \langle \mathbf{R}\boldsymbol{\xi} \rangle, \quad (\mathbf{R}\underline{\mathbf{A}}) \langle \boldsymbol{\xi} \rangle = \mathbf{R} \langle \underline{\mathbf{A}} \langle \boldsymbol{\xi} \rangle \rangle.$$

The *point product* of two states, at least one of which is scalar-valued, is defined by:

$$\begin{aligned} (\underline{a}\underline{\mathbf{A}}) \langle \boldsymbol{\xi} \rangle &= (\underline{\mathbf{A}}\underline{a}) \langle \boldsymbol{\xi} \rangle = \underline{a} \langle \boldsymbol{\xi} \rangle \underline{\mathbf{A}} \langle \boldsymbol{\xi} \rangle, \\ (\underline{a}\underline{b}) \langle \boldsymbol{\xi} \rangle &= (\underline{b}\underline{a}) \langle \boldsymbol{\xi} \rangle = \underline{a} \langle \boldsymbol{\xi} \rangle \underline{b} \langle \boldsymbol{\xi} \rangle. \end{aligned}$$

Double states operate on two bonds rather than just one and are denoted by $\underline{\mathbb{K}} \langle \xi, \eta \rangle$, where ξ and η are bonds. Double states are usually tensor-valued. An example is the dyadic product double state defined by:

$$(\underline{\mathbf{A}} \star \underline{\mathbf{B}}) \langle \xi, \eta \rangle = \underline{\mathbf{A}} \langle \xi \rangle \otimes \underline{\mathbf{B}} \langle \eta \rangle.$$

The dot product of a tensor-valued double state $\underline{\mathbb{K}}$ with a vector-valued state $\underline{\mathbf{A}}$ is the vector-valued state given by:

$$(\underline{\mathbb{K}} \bullet \underline{\mathbf{A}}) \langle \xi \rangle = \int_{\mathcal{H}_x} \underline{\mathbb{K}} \langle \xi, \eta \rangle \underline{\mathbf{A}} \langle \eta \rangle \, dV_\eta, \quad (\underline{\mathbf{A}} \bullet \underline{\mathbb{K}}) \langle \xi \rangle = \int_{\mathcal{H}_x} \underline{\mathbf{A}} \langle \eta \rangle \underline{\mathbb{K}} \langle \eta, \xi \rangle \, dV_\eta.$$

It is necessary to define a notion of differentiation of functions of states. Let $\Psi(\underline{\mathbf{A}})$ be a scalar-valued function of a vector-valued state. Consider a small increment $\Delta \underline{\mathbf{A}}$. Suppose there is a vector-valued state denoted $\Psi_{\underline{\mathbf{A}}}(\underline{\mathbf{A}})$ such that for any $\Delta \underline{\mathbf{A}}$,

$$\Psi(\underline{\mathbf{A}} + \Delta \underline{\mathbf{A}}) - \Psi(\underline{\mathbf{A}}) = \Psi_{\underline{\mathbf{A}}}(\underline{\mathbf{A}}) \bullet \Delta \underline{\mathbf{A}} + o(\|\Delta \underline{\mathbf{A}}\|). \quad (3)$$

Then $\Psi_{\underline{\mathbf{A}}}(\underline{\mathbf{A}})$ is called the *Fréchet derivative* of Ψ at $\underline{\mathbf{A}}$. The Fréchet derivative of a function whose value is a vector-valued state is a double state.

Fréchet derivatives obey a form of the chain rule. For example, if $\underline{\mathbf{B}}$ is a state-valued function of the state $\underline{\mathbf{A}}$,

$$\Psi_{\underline{\mathbf{A}}}(\underline{\mathbf{B}}(\underline{\mathbf{A}})) = \Psi_{\underline{\mathbf{B}}} \bullet \underline{\mathbf{B}}_{\underline{\mathbf{A}}},$$

or

$$\Psi_{\underline{\mathbf{A}}} \langle \xi \rangle = \int_{\mathcal{H}_x} \Psi_{\underline{\mathbf{B}}} \langle \eta \rangle \cdot \underline{\mathbf{B}}_{\underline{\mathbf{A}}} \langle \eta, \xi \rangle \, dV_\eta.$$

Note that $\underline{\mathbf{B}}_{\underline{\mathbf{A}}}$ is a tensor-valued double state. More details about the use of states can be found in Silling et al. (2007). The symmetries of double states play an interesting role in the linearized peridynamic theory, which is beyond the scope of the present chapter (Silling 2010).

Balance of Momentum

Using the nonlocal concept of strain energy described in section “Basic Concepts,” expressions for balance of linear momentum will now be derived. Assume that \mathcal{B} is bounded, and consider a time-independent deformation $\mathbf{y}(\mathbf{x})$ under external body force density field $\mathbf{b}(\mathbf{x})$. Define the total potential energy of the body by:

$$\Phi_{\mathbf{y}} = \int_{\mathcal{B}} (W(\mathbf{x}) - \mathbf{b}(\mathbf{x}) \cdot \mathbf{y}(\mathbf{x})) \, dV_{\mathbf{x}}. \quad (4)$$

Let $\Delta \mathbf{y}$ be a small increment in the deformation and assume that the Fréchet derivative $\widehat{W}_{\underline{\mathbf{Y}}}(\underline{\mathbf{Y}})$ exists for all $\underline{\mathbf{Y}}$. Define

$$\underline{\mathbf{T}}[\mathbf{x}] = \widehat{W}_{\underline{\mathbf{Y}}}(\underline{\mathbf{Y}}[\mathbf{x}]).$$

In the following, it is convenient to adopt the convention that states produce the value 0 for bonds longer than δ . This permits us to replace \mathcal{H}_x with \mathcal{B} as the region of integration in some of the expressions below. From Eqs. 2, 3, and 4,

$$\begin{aligned} \Delta \Phi_{\mathbf{y}} &= \int_{\mathcal{B}} (\underline{\mathbf{T}}[\mathbf{x}] \bullet \Delta \underline{\mathbf{Y}}[\mathbf{x}] - \mathbf{b}(\mathbf{x}) \cdot \Delta \mathbf{y}(\mathbf{x})) \, dV_{\mathbf{x}} \\ &= \int_{\mathcal{B}} \left(\int_{\mathcal{B}} \{ \underline{\mathbf{T}}[\mathbf{x}] \langle \mathbf{q} - \mathbf{x} \rangle \cdot (\Delta \mathbf{y}(\mathbf{q}) - \Delta \mathbf{y}(\mathbf{x})) \} \, dV_{\mathbf{q}} - \mathbf{b}(\mathbf{x}) \cdot \Delta \mathbf{y}(\mathbf{x}) \right) \, dV_{\mathbf{x}} \\ &= \int_{\mathcal{B}} \left(\int_{\mathcal{B}} \{ \underline{\mathbf{T}}[\mathbf{q}] \langle \mathbf{x} - \mathbf{q} \rangle \cdot \Delta \mathbf{y}(\mathbf{x}) - \underline{\mathbf{T}}[\mathbf{x}] \langle \mathbf{q} - \mathbf{x} \rangle \cdot \Delta \mathbf{y}(\mathbf{x}) \} \, dV_{\mathbf{q}} - \mathbf{b}(\mathbf{x}) \cdot \mathbf{y}(\mathbf{x}) \right) \, dV_{\mathbf{x}} \\ &= \int_{\mathcal{B}} \left(\int_{\mathcal{B}} \{ \underline{\mathbf{T}}[\mathbf{q}] \langle \mathbf{x} - \mathbf{q} \rangle - \underline{\mathbf{T}}[\mathbf{x}] \langle \mathbf{q} - \mathbf{x} \rangle \} \, dV_{\mathbf{q}} - \mathbf{b}(\mathbf{x}) \right) \cdot \Delta \mathbf{y}(\mathbf{x}) \, dV_{\mathbf{x}}, \end{aligned} \quad (5)$$

where an interchange of dummy variables of integration $\mathbf{x} \leftrightarrow \mathbf{q}$ is used in the third line. Stationary potential energy requires $\Delta \Phi_{\mathbf{y}} = 0$ for all $\Delta \mathbf{y}$ and hence Eq. 5 can be localized to yield:

$$\int_{\mathcal{H}_x} \{ \underline{\mathbf{T}}[\mathbf{x}] \langle \mathbf{q} - \mathbf{x} \rangle - \underline{\mathbf{T}}[\mathbf{q}] \langle \mathbf{x} - \mathbf{q} \rangle \} \, dV_{\mathbf{q}} + \mathbf{b}(\mathbf{x}) = 0 \quad \forall \mathbf{x} \in \mathcal{B}. \quad (6)$$

Equation 6 is the equilibrium equation in peridynamics. It is frequently written in the more suggestive form:

$$\int_{\mathcal{H}_x} \mathbf{f}(\mathbf{q}, \mathbf{x}) \, dV_{\mathbf{q}} + \mathbf{b}(\mathbf{x}) = 0 \quad \forall \mathbf{x} \in \mathcal{B} \quad (7)$$

where $\mathbf{f}(\mathbf{q}, \mathbf{x})$ is the *pairwise bond force density* field, given by:

$$\mathbf{f}(\mathbf{q}, \mathbf{x}) = \underline{\mathbf{T}}[\mathbf{x}] \langle \mathbf{q} - \mathbf{x} \rangle - \underline{\mathbf{T}}[\mathbf{q}] \langle \mathbf{x} - \mathbf{q} \rangle \quad \forall \mathbf{x}, \mathbf{q} \in \mathcal{B}.$$

Observe that $\mathbf{f}(\mathbf{q}, \mathbf{x})$ is comprised of two terms arising from the material models at \mathbf{x} and at \mathbf{q} .

The pairwise bond force density field obeys the antisymmetry condition:

$$\mathbf{f}(\mathbf{x}, \mathbf{q}) = -\mathbf{f}(\mathbf{q}, \mathbf{x}) \quad \forall \mathbf{x}, \mathbf{q} \in \mathcal{B}.$$

Mechanically, the vector $\mathbf{f}(\mathbf{q}, \mathbf{x})$ can be thought of the force density (force per unit volume squared) that \mathbf{q} exerts on \mathbf{x} . However, it is best not to take this intuitive

picture too literally, because there is not necessarily a direct physical interaction between the two points.

By d'Alembert's rule, the time-dependent equation of motion is found from Eq. 7 to be:

$$\rho(\mathbf{x}) \ddot{\mathbf{y}}(\mathbf{x}, t) = \int_{\mathcal{H}_x} \mathbf{f}(\mathbf{q}, \mathbf{x}, t) dV_{\mathbf{q}} + \mathbf{b}(\mathbf{x}, t) \quad \forall \mathbf{x} \in \mathcal{B}, t \geq 0 \quad (8)$$

where ρ is the mass density field.

A more axiomatic and general derivation of the linear momentum balance in peridynamics, which does not assume the existence of a strain-energy density field, can be found in Silling and Lehoucq (2010). Also, a derivation from statistical mechanics is available in Lehoucq and Sears (2011).

Energy Balance and Thermodynamics

The laws of thermodynamics can be written in a form compatible with the nonlocal nature of the peridynamic theory. The first law of thermodynamics takes the following form:

$$\dot{\varepsilon}(\mathbf{x}, t) = \underline{\mathbf{T}}[\mathbf{x}, t] \bullet \dot{\underline{\mathbf{Y}}}[\mathbf{x}, t] + h(\mathbf{x}, t) + s(\mathbf{x}, t) \quad \forall \mathbf{x} \in \mathcal{B}, t \geq 0,$$

where ε is the internal energy density (per unit volume), h is the rate of heat transport to \mathbf{x} , and s is the energy source rate. More simply,

$$\dot{\varepsilon} = \underline{\mathbf{T}} \bullet \dot{\underline{\mathbf{Y}}} + h + s. \quad (9)$$

To provide an intuitive picture of this expression, recall that by Eq. 2,

$$\underline{\mathbf{T}} \bullet \dot{\underline{\mathbf{Y}}} = \int_{\mathcal{H}_x} \underline{\mathbf{T}}\langle \xi \rangle \cdot \dot{\underline{\mathbf{Y}}}\langle \xi \rangle dV_{\xi}.$$

The energy balance (9) therefore sums up the rate of work done by the bond forces acting against the rate of extension of the individual bonds. The key point is that only the force state at \mathbf{x} , that is, $\underline{\mathbf{T}}[\mathbf{x}, t]$, contributes to the energy change at \mathbf{x} . The force state $\underline{\mathbf{T}}[\mathbf{q}, t]$ does not, even though it appears in the momentum balance (6). This partitioning of the pairwise bond force into energy contributions at \mathbf{x} and \mathbf{q} is the unique feature of the peridynamic version of thermodynamics that causes the internal energy density to be *additive*, provided that h is conserved. This fact apparently resolves a long-standing question about whether it is even possible to define a nonlocal internal energy density that is additive (Gurtin and Williams 1971); see (Silling and Lehoucq 2010) for a more complete discussion.

The energy balance (9) applies regardless of how the heat transport h is specified. This could be supplied by the Fourier heat conduction expression or by a nonlocal

diffusion law such as:

$$h(\mathbf{x}, t) = \int_{\mathcal{H}_x} \left[\underline{H}[\mathbf{x}, t] \langle \mathbf{q} - \mathbf{x} \rangle - \underline{H}[\mathbf{q}, t] \langle \mathbf{x} - \mathbf{q} \rangle \right] dV_q$$

where \underline{H} is a scalar-valued state. The most intuitive form of such an \underline{H} models heat diffusion along bonds as though they are conducting wires that are insulated from each other (Bobaru and Duangpanya 2010):

$$\underline{H}[\mathbf{x}, t] \langle \mathbf{q} - \mathbf{x} \rangle = \underline{K}[\mathbf{x}] \langle \mathbf{q} - \mathbf{x} \rangle (\theta(\mathbf{q}, t) - \theta(\mathbf{x}, t)) \quad (10)$$

where θ is the temperature field and K is the bond conductivity. The heat flow expression (10) is useful in modeling systems with inherently nonlocal diffusion mechanisms such as radiative heat transport, as well as in treating discontinuities and singularities in the temperature field. A model for heat transport that is more general than (Eq. 10) could have the form:

$$\underline{H}[\mathbf{x}, t] = \widehat{\underline{H}}(\underline{\Theta}[\mathbf{x}, t]), \quad \underline{\Theta}[\mathbf{x}, t] \langle \mathbf{q} - \mathbf{x} \rangle = \theta(\mathbf{q}, t) - \theta(\mathbf{x}, t). \quad (11)$$

This representation could model systems in which the heat flow in each bond can depend not only on the temperature difference between its own endpoints but also on the temperature difference in other bonds in the family.

The local form of second law of thermodynamics as implemented in peridynamics has the following form:

$$\theta \dot{\eta} \geq h + s \quad (12)$$

where η is the entropy density field (per unit volume) in the reference configuration. A restriction on the admissible forms of $\widehat{\underline{H}}$ arises from the second law, as discussed in section “[Restriction on the Heat Transport Model.](#)”

Material Models

A peridynamic material model $\widehat{\underline{\mathbf{T}}}$ determines the force state at every material point \mathbf{x} and time t . For most materials, the force state depends on the deformation state $\underline{\mathbf{Y}}[\mathbf{x}, t]$ and possibly other variables as well. If $\underline{\mathbf{Y}}$ is the only quantity that the material model depends on, we write:

$$\underline{\mathbf{T}}[\mathbf{x}, t] = \widehat{\underline{\mathbf{T}}}(\underline{\mathbf{Y}}[\mathbf{x}, t]).$$

If the body is heterogeneous, that is, if the material model depends explicitly on position, we write:

$$\underline{\mathbf{T}}[\mathbf{x}, t] = \widehat{\underline{\mathbf{T}}}(\underline{\mathbf{Y}}[\mathbf{x}, t], \mathbf{x}).$$

The material model can also depend on the rate of deformation, $\dot{\underline{\mathbf{Y}}}$, and other physically relevant quantities such as temperature (section

“Thermodynamic Form of A Material Model”), damage (section “Damage as a Thermodynamic Variable”), or plastic deformation (section “Plasticity”).

Material models are required to satisfy certain general rules. The requirement of *nonpolarity* is written as:

$$\int_{\mathcal{H}_x} \widehat{\mathbf{T}}(\mathbf{Y}) \langle \boldsymbol{\xi} \rangle \times \underline{\mathbf{Y}} \langle \boldsymbol{\xi} \rangle \, dV_{\boldsymbol{\xi}} = 0 \quad \forall \underline{\mathbf{Y}}. \quad (13)$$

Nonpolarity guarantees that global balance of angular momentum holds, that is, the material model does not nonphysically create angular momentum. It is similar to the required symmetry of the Cauchy stress tensor in the local theory. Nonpolarity is not a requirement in the micropolar versions of peridynamics that have been proposed (Gerstle et al. 2007) or in a peridynamic shell theory that includes rotational degrees of freedom (Chowdhury et al. 2016).

In the absence of external fields that provide a special physical direction that affects material response, a peridynamic material model is required to satisfy *objectivity*. This is a requirement that if the family is deformed and then rigidly rotated, then the force state undergoes the same rigid rotation. This condition is written as follows:

$$\widehat{\mathbf{T}}(\mathbf{Q}\mathbf{Y}) = \mathbf{Q}\widehat{\mathbf{T}}(\mathbf{Y}) \quad \forall \mathbf{Q}, \quad \forall \underline{\mathbf{Y}}$$

where \mathbf{Q} is any proper orthogonal tensor. Unlike objectivity, *isotropy* is not a general requirement but is appropriate for modeling materials that have no internal special direction (such as embedded unidirectional reinforcement fibers). Isotropy means that if the body is first rigidly rotated, then deformed, the force state is the same as if there were no rotation:

$$\widehat{\mathbf{T}}(\mathbf{Y}\mathbf{Q}) = \widehat{\mathbf{T}}(\mathbf{Y}) \quad \forall \mathbf{Q}, \quad \forall \underline{\mathbf{Y}}$$

where \mathbf{Q} is any proper orthogonal tensor.

If there is a strain-energy function $\widehat{W}_{\underline{\mathbf{Y}}}(\mathbf{Y})$ such that:

$$\widehat{\mathbf{T}}(\mathbf{Y}) = \widehat{W}_{\underline{\mathbf{Y}}}(\mathbf{Y}) \quad \forall \underline{\mathbf{Y}},$$

then the material is *elastic*. An elastic material is objective if and only if $\widehat{W}(\mathbf{Q}\mathbf{Y}) = \widehat{W}(\mathbf{Y}) \quad \forall \mathbf{Q}, \quad \forall \underline{\mathbf{Y}}$. An elastic material is isotropic if and only if:

$$\widehat{W}(\mathbf{Y}\mathbf{Q}) = \widehat{W}(\mathbf{Y}) \quad \forall \mathbf{Q}, \quad \forall \underline{\mathbf{Y}}.$$

A material that is elastic and objective is necessarily nonpolar (Silling 2010). This property is a convenience to developers of material models, because it is often much easier to prove objectivity of \widehat{W} than to prove nonpolarity directly from Eq. 13.

Peridynamic material models fall into the following categories:

- *Bond-based*: Each bond has a force response that depends *only* on its own deformation and damage, independent of all other bonds in the family. For a bond-based material model,

$$\underline{\mathbf{T}}[\mathbf{x}] \langle \mathbf{q} - \mathbf{x} \rangle = \widehat{\mathbf{t}}(\underline{\mathbf{Y}}[\mathbf{x}] \langle \mathbf{q} - \mathbf{x} \rangle, \mathbf{q}, \mathbf{x}).$$

Note that the function $\widehat{\mathbf{t}}$ and its arguments are all vectors rather than vector-valued states. For a homogeneous body, the equilibrium equation for a bond-based material is often written as:

$$\int_{\mathcal{H}_x} \mathbf{f}(\mathbf{u}(\mathbf{q}) - \mathbf{u}(\mathbf{x}), \mathbf{q} - \mathbf{x}) dV_{\mathbf{q}} + \mathbf{b}(\mathbf{x}) = 0,$$

with \mathbf{f} rather than $\widehat{\mathbf{T}}$ as the material model. An example is given by:

$$\mathbf{f}(\eta, \xi) = \underline{\mathbf{C}} \langle \xi \rangle (|\eta + \xi| - |\xi|) \underline{\mathbf{M}} \langle \xi \rangle, \quad \underline{\mathbf{M}} \langle \xi \rangle = \frac{\underline{\mathbf{Y}} \langle \xi \rangle}{|\underline{\mathbf{Y}} \langle \xi \rangle|}$$

where $\underline{\mathbf{C}}$ is a scalar-valued state. In this material model, each bond acts like a linear spring.

- *State based*: Each bond has a force response that can depend on the deformation and damage in *all* the bonds in the family. Among state-based material models, there are two classes: *ordinary*, in which the bond forces are always parallel to the deformed bond direction $\underline{\mathbf{M}} \langle \xi \rangle$, and all others, which are called *non-ordinary*. Ordinary materials have the advantage of automatically being nonpolar.

Thermodynamic Form of a Material Model

This section describes a thermodynamic formulation of peridynamic material models. Material response that can be defined in terms of static variables (without rates, gradients, and loading history) can often be written in terms of a free-energy function. The advantage of this way of characterizing material response is that the resulting values for the force state, temperature, entropy, damage, coefficient of thermal expansion, and heat capacity are always consistent. It also permits us to incorporate restrictions on the material model derived from the second law of thermodynamics. Additional terms such as rate dependence can be added to the force state obtained from the thermodynamic form.

Define the free-energy density at a material point \mathbf{x} and time t by:

$$\psi = \varepsilon - \theta \eta \tag{14}$$

where, as before, ε , θ , and η are the internal energy density, temperature, and entropy density, respectively. Taking the time derivative of Eq. 14 and applying the second law expression, Eq. 12 leads to:

$$\dot{\varepsilon} - \dot{\theta}\eta - \dot{\psi} \geq 0.$$

Combining this with the first law expression (9) yields:

$$\underline{\mathbf{T}} \bullet \dot{\underline{\mathbf{Y}}} + h + s - \dot{\theta}\eta - \dot{\psi} \geq 0. \tag{15}$$

Suppose the material model is expressed in terms of the dependence of free-energy density on both the deformation state and temperature, thus $\psi(\underline{\mathbf{Y}}, \theta)$. By the chain rule,

$$\dot{\psi} = \psi_{\theta}\dot{\theta} + \psi_{\underline{\mathbf{Y}}} \bullet \dot{\underline{\mathbf{Y}}}. \tag{16}$$

In this expression, ψ_{θ} is a partial derivative and $\psi_{\underline{\mathbf{Y}}}$ is a Fréchet derivative, and therefore, its value is a state. Combining Eq. 15 with Eq. 16 and regrouping terms yields:

$$(\underline{\mathbf{T}} - \psi_{\underline{\mathbf{Y}}}) \bullet \dot{\underline{\mathbf{Y}}} + h + s - (\eta + \psi_{\theta})\dot{\theta} \geq 0. \tag{17}$$

Following the reasoning of Coleman and Noll (1963; Gurtin et al. 2010), in principle, we can contrive an experiment in which $h = s = 0$ and the quantities $\dot{\underline{\mathbf{Y}}}$ and $\dot{\theta}$ are prescribed independently of each other.

Suppose that in such a thought experiment $\dot{\theta} = 0$ but $\dot{\underline{\mathbf{Y}}}$ is varied, and vice versa. Enforcing the inequality (17) then leads to the conclusions

$$\underline{\mathbf{T}} = \psi_{\underline{\mathbf{Y}}}, \quad \eta = -\psi_{\theta}. \tag{18}$$

The first of these provides the force state in the thermodynamic form of a rate-independent material model. Rate dependence can be included by assuming the dependence $\psi(\underline{\mathbf{Y}}, \dot{\underline{\mathbf{Y}}}, \theta, \dot{\theta})$, leading to an additional term in Eq. 17:

$$(\underline{\mathbf{T}} - \psi_{\underline{\mathbf{Y}}}) \bullet \dot{\underline{\mathbf{Y}}} + h + s - (\eta + \psi_{\theta})\dot{\theta} + \psi_{\dot{\theta}}\ddot{\theta} - \psi_{\dot{\underline{\mathbf{Y}}}} \bullet \ddot{\underline{\mathbf{Y}}} \geq 0.$$

From this, using the same reasoning as before, one concludes that $\psi_{\dot{\theta}} = \psi_{\dot{\underline{\mathbf{Y}}}} = 0$, that is, the free energy density cannot depend explicitly on $\dot{\theta}$ or $\dot{\underline{\mathbf{Y}}}$. However, without loss of generality, rate effects can be incorporated by partitioning the force state response into equilibrium and rate-dependent parts:

$$\widehat{\mathbf{T}}(\underline{\mathbf{Y}}, \dot{\underline{\mathbf{Y}}}, \theta) = \underline{\mathbf{T}}^e(\underline{\mathbf{Y}}, \theta) + \underline{\mathbf{T}}^d(\underline{\mathbf{Y}}, \dot{\underline{\mathbf{Y}}}, \theta), \quad \underline{\mathbf{T}}^d(\underline{\mathbf{Y}}, \underline{\mathbf{0}}, \theta) = \underline{\mathbf{0}}.$$

From this assumption, it follows by a similar Coleman-Noll type of argument (Fried 2010) that

$$\underline{\mathbf{T}}^e = \psi_{\underline{\mathbf{Y}}}, \quad \underline{\mathbf{T}}^d \bullet \dot{\underline{\mathbf{Y}}} \geq 0. \quad (19)$$

The second of Eq. 19 is a *dissipation inequality*. It implies that

$$\underline{\mathbf{T}}^d \langle \underline{\xi} \rangle \bullet \dot{\underline{\mathbf{Y}}} \langle \underline{\xi} \rangle \geq 0, \quad \forall \underline{\xi} \in \mathcal{H}_{\mathbf{x}}.$$

Work done by both $\underline{\mathbf{T}}^e$ and $\underline{\mathbf{T}}^d$ contribute to the internal energy density according to the first law (9), in general changing the temperature. These temperature changes affect ψ indirectly, even though $\psi_{\dot{\underline{\mathbf{Y}}}} = \underline{\mathbf{0}}$.

Thermal expansion arises from the coupling between $\underline{\mathbf{Y}}$ and θ in the form of ψ . For example, for an elastic material with strain-energy density function \widehat{W} , a model with thermal expansion can be defined by:

$$\psi(\underline{\mathbf{Y}}, \theta) = \frac{\widehat{W}((1 - \alpha\Delta\theta)\underline{\mathbf{Y}})}{1 - \alpha\Delta\theta}, \quad \Delta\theta := \theta - \theta_0$$

where α is a linear coefficient of thermal expansion. Then by the chain rule,

$$\underline{\mathbf{T}} = \psi_{\underline{\mathbf{Y}}} = \widehat{W}_{\underline{\mathbf{Y}}}((1 - \alpha\Delta\theta)\underline{\mathbf{Y}}) \approx \widehat{W}_{\underline{\mathbf{Y}}}(\underline{\mathbf{Y}}) - \alpha\Delta\theta \underline{\mathbb{K}} \bullet \underline{\mathbf{X}}$$

where the approximation holds if $|\alpha\Delta\theta| \ll 1$ and $|\underline{\mathbf{Y}}(\underline{\xi}) - \underline{\xi}| \ll |\underline{\xi}|$ for all $\underline{\xi}$. $\underline{\mathbb{K}}$ is the *micromodulus double state*, defined by:

$$\underline{\mathbb{K}} = \widehat{W}_{\underline{\mathbf{Y}}\underline{\mathbf{Y}}}.$$

(See Silling (2010) for details on the properties of $\underline{\mathbb{K}}$ and its role in the linearized theory.)

Restriction on the Heat Transport Model

Recall the local form of the second law of thermodynamics,

$$\theta \dot{\eta} \geq h + s. \quad (20)$$

All of us are familiar with the restriction in the local theory that heat cannot flow from cold to hot, which is a consequence of the second law. It is interesting to investigate the analogous restriction in peridynamics.

To do this, consider a bounded, nondeforming, thermodynamically isolated body. Assume that the material model is as described in the section “[Thermodynamic Form of a Material Model](#),” in which the free-energy density has the dependence $\psi(\underline{\mathbf{Y}}, \theta)$. As a consequence of the second of Eq. 18, since $\dot{\underline{\mathbf{Y}}} = \mathbf{0}$, it follows that $\dot{\eta}$ can depend only on θ . This implies that equality holds in Eq. 20, that is,

$$\theta \dot{\eta} = h. \tag{21}$$

(This is not a result of reversibility, which is not assumed here.) Now compute the total entropy change in the body. From Eqs. 11 and 21 (dropping t from the notation),

$$\begin{aligned} \int_{\mathcal{B}} \dot{\eta}(\mathbf{x}) dV_{\mathbf{x}} &= \int_{\mathcal{B}} \frac{h(\mathbf{x})}{\theta(\mathbf{x})} dV_{\mathbf{x}} \\ &= \int_{\mathcal{B}} \int_{\mathcal{B}} \frac{\underline{H}[\mathbf{x}] \langle \mathbf{q} - \mathbf{x} \rangle - \underline{H}[\mathbf{q}] \langle \mathbf{x} - \mathbf{q} \rangle}{\theta(\mathbf{x})} dV_{\mathbf{q}} dV_{\mathbf{x}} \\ &= \int_{\mathcal{B}} \int_{\mathcal{B}} \left(\frac{1}{\theta(\mathbf{x})} - \frac{1}{\theta(\mathbf{q})} \right) \underline{H}[\mathbf{x}] \langle \mathbf{q} - \mathbf{x} \rangle dV_{\mathbf{q}} dV_{\mathbf{x}} \\ &= - \int_{\mathcal{B}} \underline{\beta}[\mathbf{x}] \bullet \underline{H}[\mathbf{x}] dV_{\mathbf{x}} \end{aligned} \tag{22}$$

where $\underline{\beta}$ is the scalar-valued state defined by:

$$\underline{\beta}[\mathbf{x}] \langle \mathbf{q} - \mathbf{x} \rangle = \frac{1}{\theta(\mathbf{q})} - \frac{1}{\theta(\mathbf{x})}.$$

Suppose that the heat transport model \widehat{H} obeys

$$\underline{\beta} \bullet \widehat{H}(\Theta) \leq 0 \quad \forall \Theta, \tag{23}$$

or equivalently

$$\int_{\mathcal{H}_{\mathbf{x}}} \underline{\beta} \langle \xi \rangle \widehat{H} \langle \xi \rangle dV_{\mathbf{x}} \leq 0 \quad \forall \Theta.$$

Working backwards through the steps in Eq. 22, evidently Eq. 23 implies that

$$\int_{\mathcal{B}} \dot{\eta}(\mathbf{x}) dV_{\mathbf{x}} \geq 0,$$

which is the global form of the second law for an isolated body. Note that the stronger restriction

$$\underline{\beta} \langle \xi \rangle \widehat{H} \langle \xi \rangle \leq 0 \quad \forall \xi$$

is sufficient but not necessary for Eq. 23 to hold.

The inequality (23) is a restriction on the constitutive model for heat transport \widehat{H} ; it is the peridynamic form of the rule that “heat cannot flow from cold to hot.” Details on peridynamic modeling of heat transport can be found in Bobaru and Duangpanya (2010), Bobaru and Duangpanya (2012), and Oterkus et al. (2014b). A fully coupled thermomechanical treatment is discussed in Oterkus et al. (2014a). The relation of the restriction (Eq. 23) to the Clausius-Duhem inequality is discussed in section “[Convergence of Peridynamics to the Local Theory](#)” below.

Damage as a Thermodynamic Variable

In the previous section, it was assumed that the material model is stated in terms of the free-energy density function given by:

$$\psi(\mathbf{Y}, \theta).$$

An important characteristic of the response of real materials is that they fracture and fail. To help model this aspect of material response, it is assumed that there is a scalar-valued state field called the *damage state*, denoted by $\underline{\phi}$, that has the distinguishing feature of monotonicity over time:

$$\dot{\underline{\phi}}(\xi) \geq 0 \quad \forall \xi \in \mathcal{H}_x. \quad (24)$$

By convention, it is usually assumed that $0 \leq \underline{\phi}(\xi) \leq 1$ for all bonds ξ , with $\underline{\phi}(\xi) = 0$ representing an undamaged bond. By assuming a material model of the form:

$$\psi(\mathbf{Y}, \theta, \underline{\phi}) \quad (25)$$

and working through the free-energy inequality discussed previously, one concludes (Silling and Lehoucq 2010) that the following dissipation inequality holds:

$$\psi_{\underline{\phi}} \bullet \dot{\underline{\phi}} \leq 0, \quad (26)$$

which also implies:

$$\psi_{\underline{\phi}}(\xi) \leq 0 \quad \forall \xi \in \mathcal{H}_x. \quad (27)$$

Damage evolves according to a prescribed material-dependent *damage growth law*,

$$\dot{\underline{\phi}} = \underline{D}(\mathbf{Y}, \dot{\mathbf{Y}}, \underline{\phi})$$

where \underline{D} is a scalar state-valued function (section “[Damage Evolution](#)”).

Examples of Material Models

The following examples of material models illustrate the connection between free-energy density and mechanical forces. They also demonstrate how to evaluate Fréchet derivatives in practice.

Bond-Based Linear Material with Damage

Consider the free-energy density function defined by:

$$\psi(\underline{\mathbf{Y}}, \theta, \underline{\phi}) = \frac{1}{2} \left(\underline{\mathbf{C}}(1 - \underline{\phi}) \underline{\mathbf{e}} \right) \bullet \underline{\mathbf{e}} + c\theta (1 - \log(\theta/\theta_0)) \quad (28)$$

or equivalently

$$\psi(\underline{\mathbf{Y}}, \theta, \underline{\phi}) = \frac{1}{2} \int_{\mathcal{H}_x} \underline{\mathbf{C}}(\xi) (1 - \underline{\phi}(\xi)) \underline{\mathbf{e}}^2(\xi) dV_\xi + c\theta (1 - \log(\theta/\theta_0)) \quad (29)$$

where $\underline{\mathbf{C}}$ is the prescribed scalar-valued *micromodulus state* and c is the heat capacity (at constant $\underline{\mathbf{Y}}$). The scalar-valued state $\underline{\mathbf{e}}$ is the *extension state*, defined by:

$$\underline{\mathbf{e}}(\xi) = |\underline{\mathbf{Y}}(\xi)| - |\xi|. \quad (30)$$

To explain how to evaluate the force state $\underline{\mathbf{T}} = \psi_{\underline{\mathbf{Y}}}$, we go through in detail the process of obtaining the Fréchet derivative. Recalling (Eq. 3), we seek to express a first-order approximation for incremental changes in ψ in response to any small $\Delta\underline{\mathbf{Y}}$ in the form:

$$\Delta\psi = \int_{\mathcal{H}_x} [\text{something}] \cdot \Delta\underline{\mathbf{Y}}(\xi) dV_\xi.$$

The process of finding the Fréchet derivative consists of finding an expression of this form. The [something], which is a function of the dummy variable of integration ξ (and therefore is a state), is the Fréchet derivative. For the example material model (29),

$$\Delta\psi = \underline{\mathbf{T}} \bullet \Delta\underline{\mathbf{Y}} = \int_{\mathcal{H}_x} \underline{\mathbf{C}}(\xi) (1 - \underline{\phi}(\xi)) \underline{\mathbf{e}}(\xi) \Delta\underline{\mathbf{e}}(\xi) dV_\xi.$$

Since

$$\Delta\underline{\mathbf{e}}(\xi) = \underline{\mathbf{M}}(\xi) \cdot \Delta\underline{\mathbf{Y}}(\xi), \quad \underline{\mathbf{M}}(\xi) = \frac{\underline{\mathbf{Y}}(\xi)}{|\underline{\mathbf{Y}}(\xi)|},$$

it follows that

$$\underline{\mathbf{T}} \bullet \Delta \underline{\mathbf{Y}} = \int_{\mathcal{H}_x} \left[\underline{C} \langle \xi \rangle \left(1 - \underline{\phi} \langle \xi \rangle \right) \underline{e} \langle \xi \rangle \underline{\mathbf{M}} \langle \xi \rangle \right] \cdot \Delta \underline{\mathbf{Y}} \langle \xi \rangle dV_{\xi}.$$

Comparing this with Eq. 3 leads to the conclusion that

$$\widehat{\underline{\mathbf{T}}}(\underline{\mathbf{Y}}) = \underline{C} \left(\underline{\mathbf{1}} - \underline{\phi} \right) \underline{e} \underline{\mathbf{M}}. \quad (31)$$

or

$$\underline{\mathbf{T}} \langle \xi \rangle = \underline{C} \langle \xi \rangle \left(\underline{\mathbf{1}} - \underline{\phi} \langle \xi \rangle \right) \underline{e} \langle \xi \rangle \underline{\mathbf{M}} \langle \xi \rangle \quad \forall \xi \in \mathcal{H}_x.$$

Some features of the material model (31) are as follows:

- The bond force density vector in each bond is parallel to the deformed bond.
- Holding damage fixed, the magnitude of the bond force density varies linearly with bond extension.
- The model is *bond-based*: each bond responds independently of all the others.
- The model is geometrically nonlinear: it allows for large deformation.
- Increasing the damage in each bond decreases the magnitude of each bond force density in that bond.

From Eq. 28 and the second of Eq. 18, one finds that for this material model,

$$\eta = -\psi_{\theta} = c \log(\theta/\theta_0) \quad (32)$$

From Eq. 14, the second of Eq. 18, and the time derivative of Eq. 32, it follows that if $\underline{\mathbf{Y}}$ and $\underline{\phi}$ are held constant,

$$\begin{aligned} \dot{\epsilon} &= \psi_{\theta} \dot{\theta} + \dot{\theta} \eta + \theta \dot{\eta} \\ &= \theta \dot{\eta} \\ &= c \dot{\theta}. \end{aligned}$$

Thus, unlike the free energy, the internal energy varies linearly with temperature under these conditions, with the heat capacity as the constant of proportionality.

Ordinary State-Based Linear Material with Damage

A modification of Eq. 31 that includes a volume change may be written as:

$$\widehat{\underline{\mathbf{T}}}(\underline{\mathbf{Y}}) = \left(\underline{C} \left(\underline{\mathbf{1}} - \underline{\phi} \right) \underline{e} + A \vartheta \underline{\mathbf{1}} \right) \underline{\mathbf{M}}, \quad \vartheta = \frac{\underline{\mathbf{1}} \bullet \underline{e}}{\underline{\mathbf{1}} \bullet \underline{\mathbf{1}}}$$

where A is a constant and ϑ is the dilatation. (See Eq. 40 below for a more general definition of dilatation.) Note that the dilatation depends on all the bonds in the family.

Non-ordinary State-Based Material

The following material model tends to resist bending:

$$\widehat{\mathbf{T}}(\underline{\mathbf{Y}})(\xi) = \begin{cases} A(\underline{e}(\xi) + \underline{e}(-\xi))(\underline{\mathbf{Y}}(\xi) - \underline{\mathbf{Y}}(-\xi)) & \text{if } \xi \in \mathcal{H}_x \text{ and } -\xi \in \mathcal{H}_x, \\ 0 & \text{otherwise.} \end{cases}$$

where A is a constant. Another way to write this is to define the *reversal state* by

$$\underline{\mathbf{R}}(\xi) = -\xi \quad \forall \xi \in \mathcal{H}_x,$$

then

$$\widehat{\mathbf{T}}(\underline{\mathbf{Y}}) = A(\underline{e} + \underline{e} \circ \underline{\mathbf{R}})(\underline{\mathbf{Y}} - \underline{\mathbf{Y}} \circ \underline{\mathbf{R}}).$$

This material is nonpolar even though the bond forces are not necessarily parallel to the deformed bonds. It is interesting that a state-based peridynamic material model *without changing the equilibrium equation* can resist bending; see Diyaroglu et al. (2015) and Grady and Foster (2014), for details and specific material models. This is in contrast to the standard theory of beams and plates, in which a special fourth-order PDE replaces the fundamental second-order PDEs of local continuum mechanics.

Bond-Based Viscoelastic Material

The material model (31) can be modified to include a rate-dependent damping term:

$$\widehat{\mathbf{T}}(\underline{\mathbf{Y}}, \dot{\underline{\mathbf{Y}}}) = \left(\underline{\mathbf{C}}(1 - \underline{\phi}) \underline{e} + A \dot{\underline{e}} \right) \underline{\mathbf{M}}$$

where A is a nonnegative constant. More information on viscoelastic peridynamic models can be found in Mitchell (2011a) and Weckner and Mohamed (2013).

Isotropic Bond-Based Material

In the material model (31), different bonds can have different stiffness, because $\underline{\mathbf{C}}$ is a state (that is, its value depends on the bond). (See Hu et al. 2012a; Oterkus and

Madenci 2012). for examples of anisotropic material models.) To model isotropic materials, set

$$\underline{C}(\xi) = C^i(|\xi|)$$

where C^i is a function of bond length only, thus

$$\underline{\mathbf{T}} = \left(C^i \left(\underline{1} - \underline{\phi} \right) \underline{e} \right) \underline{\mathbf{M}}.$$

The most general form for C^i is discussed in Silling (2000).

Nonconvex Bond-Based Material

If the bond force density is not a monotonic function of extension, and if it is elastic, then its strain-energy density function is called *nonconvex*. An example is:

$$\underline{\mathbf{T}}(\xi) = A \underline{e}(\xi) \exp(-\underline{e}^2(\xi)) \underline{\mathbf{M}}(\xi) \text{ or } \widehat{\underline{\mathbf{T}}}(\underline{\mathbf{Y}}) = A \underline{e} \exp(-\underline{e}^2) \underline{\mathbf{M}}.$$

The properties of nonconvex bond-based materials, including their stability and relation to brittle fracture, are discussed by Lipton (2014, 2016). More general concepts of convexity that apply to state-based materials are, to the best of the author's knowledge, a totally unexplored area.

Discrete Particles as Peridynamic Materials

Consider a set of N particles with equal mass m that interact through a multibody potential such that the potential energy of particle i is given by:

$$\Phi_i = U_i(\mathbf{r}_{1i}, \mathbf{r}_{2i}, \dots, \mathbf{r}_{Ni}) - \mathbf{b}_i \cdot \mathbf{y}_i, \quad \mathbf{r}_{ji} := \mathbf{y}_j - \mathbf{y}_i$$

where \mathbf{y}_i and \mathbf{b}_i are the position of and external force on particle i . The potential energy of the entire set of particles is found from:

$$\Phi = \sum_{i=1}^N \Phi_i.$$

The acceleration of each particle i is obtained from Newton's second law in the form:

$$\begin{aligned} \mathbf{F}_i &= m \ddot{\mathbf{y}}_i = - \frac{\partial \Phi}{\partial \mathbf{y}_i} \\ &= \sum_{j=1}^N \left(\frac{\partial U_i}{\partial \mathbf{r}_{ji}} - \frac{\partial U_j}{\partial \mathbf{r}_{ij}} \right) + \mathbf{b}_i. \end{aligned} \quad (33)$$

Now consider the peridynamic body with mass-density field and body-force density field defined by:

$$\rho(\mathbf{x}) = \sum_{i=1}^N m \Delta(\mathbf{x} - \mathbf{x}_i), \quad \mathbf{b}(\mathbf{x}) = \sum_{i=1}^N \mathbf{b}_i \Delta(\mathbf{x} - \mathbf{x}_i) \tag{34}$$

where $\Delta(\cdot)$ denotes the 3D Dirac delta function. Here, the reference positions of the particles \mathbf{x}_i are arbitrary and merely serve to identify the particles for purposes of the mathematics. For this peridynamic body, let the material model be elastic with strain-energy density function given by:

$$\widehat{W}(\underline{\mathbf{Y}}[\mathbf{x}], \mathbf{x}) = \sum_{i=1}^N U_i(\mathbf{r}_{1i}, \mathbf{r}_{2i}, \dots, \mathbf{r}_{Ni}) \Delta(\mathbf{x} - \mathbf{x}_i), \quad \mathbf{r}_{ji} = \underline{\mathbf{Y}}[\mathbf{x}_i] \langle \mathbf{x}_j - \mathbf{x}_i \rangle. \tag{35}$$

After evaluating the Fréchet derivative of this \widehat{W} , the force state field is found to be:

$$\underline{\mathbf{T}}[\mathbf{x}] \langle \mathbf{q} - \mathbf{x} \rangle = \sum_{i=1}^N \sum_{j=1}^N \frac{\partial U_i}{\partial \mathbf{r}_{ji}} \Delta(\mathbf{x} - \mathbf{x}_i) \Delta(\mathbf{q} - \mathbf{x}_j). \tag{36}$$

Evaluating the acceleration field using Eqs. 8, 34, and 36, the terms involving $\Delta(\mathbf{q} - \mathbf{x}_j)$ integrate to 1 and hence

$$\rho(\mathbf{x}) \ddot{\mathbf{y}}(\mathbf{x}) = \int_B \left\{ \underline{\mathbf{T}}[\mathbf{x}] \langle \mathbf{q} - \mathbf{x} \rangle - \underline{\mathbf{T}}[\mathbf{q}] \langle \mathbf{x} - \mathbf{q} \rangle \right\} dV_{\mathbf{q}} + \mathbf{b}(\mathbf{x})$$

becomes

$$\begin{aligned} \sum_{i=1}^N (m \ddot{\mathbf{y}}(\mathbf{x}) - \mathbf{b}(\mathbf{x})) \Delta(\mathbf{x} - \mathbf{x}_i) &= \sum_{i=1}^N \sum_{j=1}^N \frac{\partial U_i}{\partial \mathbf{r}_{ji}} (\Delta(\mathbf{x} - \mathbf{x}_i) - \Delta(\mathbf{x} - \mathbf{x}_j)) \\ &= \sum_{i=1}^N \sum_{j=1}^N \left(\frac{\partial U_i}{\partial \mathbf{r}_{ji}} - \frac{\partial U_j}{\partial \mathbf{r}_{ij}} \right) \Delta(\mathbf{x} - \mathbf{x}_i) \end{aligned}$$

which implies

$$m \ddot{\mathbf{y}}(\mathbf{x}_i) = \sum_{j=1}^N \left(\frac{\partial U_i}{\partial \mathbf{r}_{ji}} - \frac{\partial U_j}{\partial \mathbf{r}_{ij}} \right) + \mathbf{b}_i, \quad i = 1, 2, \dots, N.$$

So, the peridynamic equation of motion for the body specified in Eqs. 34 and 35 reduces to Newton’s second law, (Eq. 33).

The applicability of peridynamics to both continuous and discrete systems can be useful in modeling the interaction of particles with continuous bodies. For example, certain aspects of the mechanics of a suspension of particles in a liquid can be treated simply by adding the responses of the two media, one discrete and one continuous:

$$\widehat{W}(\underline{\mathbf{Y}}, \mathbf{x}) = \sum_{i=1}^N U_i(\underline{\mathbf{Y}}, \mathbf{x}) \Delta(\mathbf{x} - \mathbf{x}_i) + \widehat{W}_c(\underline{\mathbf{Y}})$$

where \widehat{W}_c is the strain-energy density function for the continuum. Since the machinery of peridynamics can be applied to this unconventional medium, this model could potentially be used to study interesting phenomena such as wave dispersion, attenuation, scattering in suspensions of interacting particles.

Effectively Eulerian Material Models

The material models described up to now in this chapter have been Lagrangian; they refer explicitly to a reference configuration, and the bond forces arise from movement of the bonds from their reference positions. For modeling fluids under large deformations, the Lagrangian approach becomes impractical because of the gross distortion of the families. In these cases, an Eulerian approach to material modeling may be preferable.

An effectively Eulerian material model for a fluid can be derived by letting the horizon be infinite but limiting the response to bonds that *currently* have length less than a prescribed distance δ in the deformed configuration. For example, such a model for a fluid could be specified by defining a *nonlocal density* as follows:

$$\bar{\rho}(\mathbf{x}) = \rho_0 \int_{\mathcal{B}} \omega(|\underline{\mathbf{Y}}[\mathbf{x]} \langle \mathbf{q} - \mathbf{x} \rangle|) \mathbf{d}V_{\xi} \quad (37)$$

where ρ_0 is the reference density and ω is a differentiable weighting function on $[0, \infty)$ such that:

$$\int_{\mathcal{B}} \omega(|\xi|) \mathbf{d}V_{\xi} = 1, \\ \omega = 0 \text{ on } [\delta, \infty), \quad \omega' \leq 0 \text{ on } [0, \delta].$$

Even though the region of integration in Eq. 37 is \mathcal{B} in the reference configuration, in effect only a neighborhood of radius δ in the deformed configuration needs to be computed. The nonlocal density Eq. 37 can be used in any conventional equation of state. The energy balance (9) continues to apply without change, since its form is independent of the material model. The pressure from the equation of state determines the bond forces through the usual Fréchet derivative:

$$\begin{aligned} \underline{\mathbf{T}}(\xi) &= \widehat{\underline{\mathbf{T}}}(\underline{\mathbf{Y}})(\xi) = \psi_{\underline{\mathbf{Y}}} = \frac{\partial \psi}{\partial v} \frac{\partial v}{\partial \bar{\rho}} \bar{\rho}_{\underline{\mathbf{Y}}} \\ &= (-p) \left(\frac{-\rho_0}{\bar{\rho}^2} \right) (\rho_0 \omega'(|\underline{\mathbf{Y}}(\xi)|) \underline{\mathbf{M}}(\xi)) \\ &= p v^2 \omega'(|\underline{\mathbf{Y}}(\xi)|) \underline{\mathbf{M}}(\xi) \quad \forall \xi \in \mathcal{B} \end{aligned}$$

where p is the pressure and $v = \rho_0/\bar{\rho}$ is the nonlocal relative volume. This approach to modeling fluids has been successfully applied to very large deformations and strong shock waves (Silling et al. 2017).

Some Lagrangian material models can be converted to effectively Eulerian models. To do this, we must eliminate any explicit dependence of $\underline{\mathbf{T}}(\xi)$ on ξ except as an identifier for bonds. For example, recall the example bond-based material model (30), (31) with $\phi = \underline{\mathbf{0}}$:

$$\underline{\mathbf{T}}(\xi) = \widehat{\underline{\mathbf{T}}}(\underline{\mathbf{Y}})(\xi) = \underline{\mathbf{C}}(\xi) \underline{\mathbf{e}}(\xi) \underline{\mathbf{M}}(\xi), \quad \underline{\mathbf{e}}(\xi) = |\underline{\mathbf{Y}}(\xi)| - |\xi|.$$

This model contains ξ explicitly through the $|\xi|$ term, so it is Lagrangian. But consider this alternative model:

$$\widehat{\underline{\mathbf{T}}}(\underline{\mathbf{Y}}, \dot{\underline{\mathbf{Y}}})(\xi) := \underline{\mathbf{C}}(\underline{\mathbf{Y}}(\xi)) \underline{\mathbf{E}}(\xi) \underline{\mathbf{M}}(\xi) \tag{38}$$

where $\underline{\mathbf{E}}$ is a scalar-valued state, that is a function of time, defined by:

$$\underline{\mathbf{E}}[0] = \underline{\mathbf{0}}, \quad \dot{\underline{\mathbf{E}}}[t] = \begin{cases} \dot{\underline{\mathbf{Y}}}(\xi) \cdot \underline{\mathbf{M}}(\xi) & \text{if } |\underline{\mathbf{Y}}(\xi)| \leq \delta, \\ 0 & \text{otherwise,} \end{cases}$$

or, more succinctly,

$$\widehat{\underline{\mathbf{T}}}(\underline{\mathbf{Y}}, \dot{\underline{\mathbf{Y}}}) = (\underline{\mathbf{C}} \circ \underline{\mathbf{Y}}) \underline{\mathbf{E}} \underline{\mathbf{M}}, \quad \dot{\underline{\mathbf{E}}} = \dot{\underline{\mathbf{Y}}} \cdot \underline{\mathbf{M}}. \tag{39}$$

Since ξ does not appear explicitly in Eq. 38 except as an identifier, this alternative model is effectively Eulerian. In Eq. 39, it is assumed that $\underline{\mathbf{C}}(\mathbf{p}) = 0$ whenever $|\mathbf{p}| > \delta$. Interactions can occur in bonds that start out with length greater than δ but get shorter over time. Similarly, bonds that are initially short will have zero bond force density if they elongate over time to length greater than δ . The material model (39) is effectively Eulerian but is not elastic.

Plasticity

For small deformations, it is conventional to express volume changes in the form of the *dilatation*, denoted ϑ . By linearization of Eq. 37, this is found to be:

$$v - 1 \approx \vartheta(\underline{\mathbf{Y}}) := \frac{3(\underline{\Omega x}) \bullet \underline{\mathbf{e}}}{(\underline{\Omega x}) \bullet \underline{\mathbf{x}}}, \quad \underline{\Omega}(\xi) = -|\xi| \omega'(|\xi|), \quad \underline{\mathbf{x}}(\xi) = |\xi|. \tag{40}$$

where \underline{e} is the extension state defined in Eq. 30. In terms of rates, since $\dot{\underline{e}} = \underline{\mathbf{M}} \cdot \dot{\underline{\mathbf{Y}}}$, Eq. 40 can be written as

$$\dot{\psi}(\dot{\underline{\mathbf{Y}}}) := \frac{3(\underline{\Omega}\underline{\mathbf{X}}) \bullet \dot{\underline{\mathbf{Y}}}}{(\underline{\Omega}\underline{\mathbf{X}}) \bullet \underline{\mathbf{X}}}. \quad (41)$$

We will not make further use of the connection between $\underline{\Omega}$ and ω given in the second of Eq. 40, so $\underline{\Omega}$ can be regarded as essentially arbitrary, except that it must be nonnegative and depend only on $|\underline{\xi}|$. Similarly, for small deformations, the pressure is given by:

$$p(\underline{\mathbf{T}}) = -\frac{1}{3}\underline{\mathbf{X}} \bullet \underline{\mathbf{T}}. \quad (42)$$

Plastic deformation can be incorporated into an elastic material model by introducing a new vector-valued state called the *permanent deformation state*, denoted $\underline{\mathbf{P}}$. Given a free-energy density function $\psi^0(\underline{\mathbf{Y}}, \theta)$, define a new free-energy density function by

$$\psi(\underline{\mathbf{Y}}, \underline{\mathbf{P}}, \theta) = \psi^0(\underline{\mathbf{Y}} - \underline{\mathbf{P}}, \theta).$$

Since $\underline{\mathbf{T}} = \psi_{\underline{\mathbf{Y}}}$, it follows that

$$\widehat{\underline{\mathbf{T}}}(\underline{\mathbf{Y}}, \underline{\mathbf{P}}, \theta) = \widehat{\underline{\mathbf{T}}}^0(\underline{\mathbf{Y}} - \underline{\mathbf{P}}, \theta)$$

and similarly

$$\psi_{\underline{\mathbf{P}}} = -\underline{\mathbf{T}}. \quad (43)$$

By repeating the steps leading up to Eq. 17, one finds that

$$(\underline{\mathbf{T}} - \psi_{\underline{\mathbf{Y}}}) \bullet \dot{\underline{\mathbf{Y}}} - \psi_{\underline{\mathbf{P}}} \bullet \dot{\underline{\mathbf{P}}} + h + s - (\eta + \psi_{\theta}) \dot{\theta} \geq 0. \quad (44)$$

From Eqs. 43 and 44, it follows that

$$\dot{\psi}^p := \underline{\mathbf{T}} \bullet \dot{\underline{\mathbf{P}}} \geq 0 \quad (45)$$

where $\dot{\psi}^p$ is the rate of *plastic work*. Equation 45 is the dissipation inequality for plastic materials.

Plastic flow can occur when the force state is on or outside of a *yield surface* defined by:

$$\mathcal{P}(\underline{\mathbf{T}}) = 0$$

where \mathcal{P} is a scalar-valued function. A possible evolution law for $\underline{\mathbf{P}}$ is given by:

$$\dot{\underline{\mathbf{P}}} = \lambda \mathcal{P}_{\underline{\mathbf{T}}}, \tag{46}$$

where $\lambda > 0$. Equation 46 can be thought of as an *associated* flow rule. The dissipation inequality (Eq. 45), which is a consequence of the second law of thermodynamics, places a restriction on \mathcal{P} :

$$\mathcal{P}_{\underline{\mathbf{T}}} \bullet \underline{\mathbf{T}} \geq 0 \quad \forall \underline{\mathbf{T}},$$

which is a type of convexity condition on the yield surface.

Many materials, especially metals under moderate stress, have yield surfaces that are nearly independent of the pressure. To account for this, the peridynamic yield surface can be defined to be a function of the *deviatoric force state*, which is obtained by subtracting off the hydrostatic part of the force state:

$$\underline{\mathbf{T}}^d = \underline{\mathbf{T}} - \frac{\underline{\mathbf{X}} \bullet \underline{\mathbf{T}}}{(\underline{\Omega} \underline{\mathbf{X}}) \bullet \underline{\mathbf{X}}} \underline{\Omega} \underline{\mathbf{X}}. \tag{47}$$

From Eqs. 42 and 47,

$$p(\underline{\mathbf{T}}^d) = 0,$$

that is, the deviatoric force state has zero pressure. For a material model in which the yield surface depends only on $\underline{\mathbf{T}}^d$, the associated flow rule (Eq. 46) can be evaluated using the chain rule for Fréchet derivatives, with the result:

$$\begin{aligned} \dot{\underline{\mathbf{P}}} &= \lambda \mathcal{P}_{\underline{\mathbf{T}}} \\ &= \lambda \mathcal{P}_{\underline{\mathbf{T}}^d} \bullet \underline{\mathbf{T}}^d \\ &= \lambda \left(\mathcal{P}_{\underline{\mathbf{T}}^d} - \frac{(\underline{\Omega} \underline{\mathbf{X}}) \bullet \mathcal{P}_{\underline{\mathbf{T}}^d}}{(\underline{\Omega} \underline{\mathbf{X}}) \bullet \underline{\mathbf{X}}} \underline{\mathbf{X}} \right). \end{aligned} \tag{48}$$

Comparing the structure of Eqs. 47 and 48, it is subtle but significant that in the latter, the $\underline{\Omega}$ is shifted from outside the fraction to inside the numerator. Using this fact, Eqs. 41 and 48 imply that:

$$\dot{\vartheta}(\dot{\underline{\mathbf{P}}}) = 0,$$

that is, the associated flow rule applied to a yield surface that depends only on the deviatoric force state results in zero volume change. This echoes the familiar result in the plasticity of metals that plastic strain has zero dilatation. Additional details on modeling plasticity within peridynamics can be found in Foster et al. (2010),

Madenci and Oterkus (2016), Mitchell (2011b), Sun and Sundararaghavan (2014), and Warren et al. (2009).

Damage Evolution

The monotonicity condition (24) is the only general requirement on the evolution of damage:

$$\dot{\underline{\phi}} \langle \underline{\xi} \rangle = \underline{D} \langle \underline{\xi} \rangle \geq 0 \quad \forall \underline{\xi} \in \mathcal{H}_x.$$

Otherwise, we are free to dream up damage growth laws. The second law restriction (Eq. 26) is really a condition on the material model, not the damage growth law.

One approach to specifying how damage grows is to define a failure surface in state space:

$$\mathcal{S}(\underline{\mathbf{T}}, \underline{\mathbf{Y}}, \underline{\phi}) = 0$$

such that damage does not increase if \mathcal{S} is in the interior of the surface, that is, if $\mathcal{S} < 0$.

An example of a plausible damage growth law is given by:

$$\underline{D} \langle \underline{\xi} \rangle = \begin{cases} 0 & \text{if } \mathcal{S} < 0, \\ \lambda \underline{F} \langle \underline{\xi} \rangle & \text{otherwise.} \end{cases} \quad (49)$$

where $\lambda(t)$ is a nonnegative scalar-valued function and \underline{F} is the *thermodynamic force state* defined by:

$$\underline{F} = -\psi_{\underline{\phi}}. \quad (50)$$

The damage growth law (49) satisfies the monotonicity condition (24) because of the result (27), which is a consequence of the dissipation inequality for damage:

$$\dot{\underline{\phi}} \langle \underline{\xi} \rangle = \underline{D} \langle \underline{\xi} \rangle = \lambda \underline{F} \langle \underline{\xi} \rangle = -\lambda \psi_{\underline{\phi}} \langle \underline{\xi} \rangle \geq 0 \quad \forall \underline{\xi} \in \mathcal{H}_x.$$

The remaining question is how to determine λ .

In dynamics, we can reasonably assume a dependence of the form:

$$\lambda = a\mathcal{S} \quad (51)$$

where a is a nonnegative constant. This relation allows the value of \mathcal{S} to be outside the failure surface, that is, $\mathcal{S} > 0$, while damage is evolving. It is interesting to investigate the stability of the resulting damage growth. Consider a uniformly

deformed body and hold the deformation fixed. Allow damage to evolve and analyze whether damage growth speeds up or slows down over time. To study this, take the total time derivative of $S(\underline{\mathbf{T}}, \underline{\mathbf{Y}}, \underline{\phi})$ holding $\dot{\underline{\mathbf{Y}}} = 0$:

$$\dot{S} = (\mathcal{S}_{\underline{\mathbf{T}}} \bullet \underline{\mathbf{T}}_{\underline{\phi}} + \mathcal{S}_{\underline{\phi}}) \bullet \dot{\underline{\phi}}.$$

Using Eqs. 49 and 50 yields:

$$\dot{\lambda} = a\dot{S} = a\lambda (\mathcal{S}_{\underline{\mathbf{T}}} \bullet \underline{\mathbf{T}}_{\underline{\phi}} + \mathcal{S}_{\underline{\phi}}) \bullet \underline{F}.$$

The solution to this ODE for λ is given by:

$$\lambda(t) = \lambda(0)e^{art}, \quad r = (\mathcal{S}_{\underline{\mathbf{T}}} \bullet \underline{\mathbf{T}}_{\underline{\phi}} + \mathcal{S}_{\underline{\phi}}) \bullet \underline{F}. \quad (52)$$

Thus, the damage growth is stable if $r \leq 0$ and unstable otherwise. Mechanically, this criterion says that if S is being driven toward the failure surface by the $\mathcal{S}_{\underline{\mathbf{T}}}$ term faster than it is being pushed away from it by the $\mathcal{S}_{\underline{\phi}}$ term (if this term is positive), then it is unstable.

In the case of quasi-static deformation, it can be assumed that, instead of Eq. 51, the condition that determines the growth of damage is that the system always remains on the failure surface as $\underline{\mathbf{Y}}$ changes, thus:

$$\dot{S} = 0.$$

Under this assumption, writing out the time derivative of S using Eq. 49 yields:

$$\begin{aligned} 0 &= \dot{S}(\underline{\mathbf{T}}, \underline{\mathbf{Y}}, \underline{\phi}) \\ &= \mathcal{S}_{\underline{\mathbf{T}}} \bullet (\underline{\mathbf{T}}_{\underline{\mathbf{Y}}} \bullet \dot{\underline{\mathbf{Y}}} + \underline{\mathbf{T}}_{\underline{\phi}} \bullet \dot{\underline{\phi}}) + \mathcal{S}_{\underline{\mathbf{Y}}} \bullet \dot{\underline{\mathbf{Y}}} + \mathcal{S}_{\underline{\phi}} \bullet \dot{\underline{\phi}} \\ &= \mathcal{S}_{\underline{\mathbf{T}}} \bullet (\underline{\mathbf{T}}_{\underline{\mathbf{Y}}} \bullet \dot{\underline{\mathbf{Y}}} + \lambda \underline{\mathbf{T}}_{\underline{\phi}} \bullet \underline{F}) + \mathcal{S}_{\underline{\mathbf{Y}}} \bullet \dot{\underline{\mathbf{Y}}} + \lambda \mathcal{S}_{\underline{\phi}} \bullet \underline{F}. \end{aligned}$$

Solving this for λ and applying the monotonicity requirement for damage leads to:

$$\lambda = \max \left\{ 0, -\frac{(\mathcal{S}_{\underline{\mathbf{T}}} \bullet \underline{\mathbf{T}}_{\underline{\mathbf{Y}}} + \mathcal{S}_{\underline{\mathbf{Y}}}) \bullet \dot{\underline{\mathbf{Y}}}}{r} \right\} \forall \dot{\underline{\mathbf{Y}}} \quad (53)$$

where r is given by the second of Eq. 52. Observe that Eq. 53 blows up as $r \rightarrow 0$, indicating the onset of unstable damage growth, as discussed previously. The relations (49) and (53) allow us to explicitly determine the rate of damage growth for every bond at every point in the body, provided the deformation is quasi-static and $S = 0$:

$$\dot{\underline{\phi}} = \max \left\{ 0, -\frac{(\mathcal{S}_{\underline{\mathbf{T}}} \bullet \underline{\mathbf{T}}_{\underline{\mathbf{Y}}} + \mathcal{S}_{\underline{\mathbf{Y}}}) \bullet \dot{\underline{\mathbf{Y}}}}{(\mathcal{S}_{\underline{\mathbf{T}}} \bullet \underline{\mathbf{T}}_{\underline{\phi}} + \mathcal{S}_{\underline{\phi}}) \bullet \underline{F}} \right\} \underline{F} \quad \forall \dot{\underline{\mathbf{Y}}}. \quad (54)$$

As an example of a failure surface, consider the material model (28) with a slight modification that introduces a binary variable to indicate intact bonds or *broken* bonds:

$$\underline{b} \langle \underline{\xi} \rangle = \begin{cases} 1 & \text{if } \underline{\phi} \langle \underline{\xi} \rangle < 1, \\ 0 & \text{otherwise.} \end{cases}$$

The modified free-energy expression is given by:

$$\psi(\underline{\mathbf{Y}}, \theta, \underline{\phi}) = \frac{1}{2} (\underline{C} (1 - \underline{\phi}) \underline{b} \underline{e}) \bullet \underline{e} + c\theta (1 - \log(\theta/\theta_0)).$$

An example of a failure surface is given by:

$$\mathcal{S}(\underline{\mathbf{T}}, \underline{\phi}) = (\underline{k}\underline{\mathbf{T}}) \bullet \underline{\mathbf{M}} + \frac{s_1}{2} \|\underline{\phi}\|^2 - s_0 \quad (55)$$

where s_0 and s_1 are constants, $s_0 \geq 0$, and \underline{k} is a nonnegative, constant, scalar-valued state. The bond force density is given by:

$$\underline{\mathbf{T}} = \psi_{\underline{\mathbf{Y}}} = \underline{C} (1 - \underline{\phi}) \underline{b} \underline{e} \underline{\mathbf{M}}.$$

Evaluating the required Fréchet derivatives in Eq. 54 leads to:

$$\begin{aligned} \underline{F} &= -\psi_{\underline{\phi}} = \frac{1}{2} \underline{C} \underline{b} \underline{e}^2, & \underline{\mathbf{T}}_{\underline{\mathbf{Y}}} &= \underline{C} (1 - \underline{\phi}) \underline{b} \underline{\mathbf{M}} \star \underline{\mathbf{M}}, & \underline{\mathbf{T}}_{\underline{\phi}} &= -\underline{C} \underline{b} \underline{e} \underline{\mathbf{M}}, \\ \mathcal{S}_{\underline{\mathbf{Y}}} &= \underline{0}, & \mathcal{S}_{\underline{\mathbf{T}}} &= \underline{k} \underline{\mathbf{M}}, & \mathcal{S}_{\underline{\phi}} &= s_1 \|\underline{\phi}\| \underline{\phi}. \end{aligned}$$

Suppose that a specimen has zero deformation and damage at time 0 and then is deformed homogeneously and quasi-statically. By Eqs. 49 and 55, damage first starts growing when the condition $\mathcal{S} = 0$ occurs, hence:

$$(\underline{k}\underline{\mathbf{T}}) \bullet \underline{\mathbf{M}} = s_0.$$

Under continued quasi-static deformation with stable damage growth, Eq. 54 then leads to:

$$\dot{\underline{\phi}} = \underline{D} = \frac{(\underline{k}\underline{C} (1 - \underline{\phi}) \underline{b}) \bullet \dot{\underline{e}}}{(\underline{k}\underline{C} \underline{b} \underline{e} - s_1 \|\underline{\phi}\| \underline{\phi}) \bullet (\underline{C} \underline{b} \underline{e}^2)} \underline{C} \underline{b} \underline{e}^2. \quad (56)$$

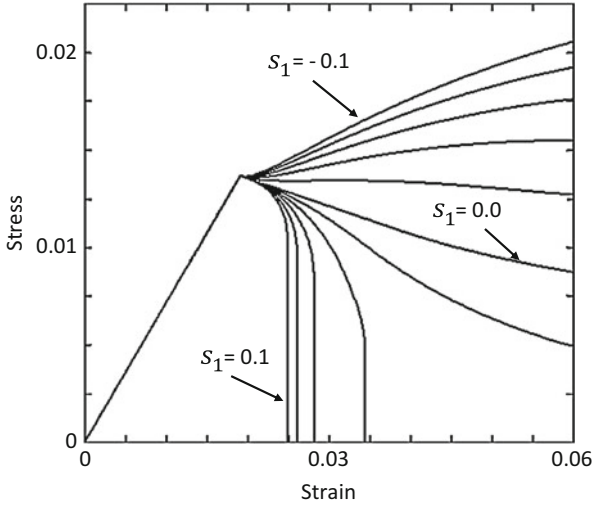


Fig. 2 Stress-strain curves for a bar under tension

Damage growth is stable until the denominator in Eq. 56 becomes nonpositive; then it becomes unstable. Figure 2 illustrates the behavior of this example material and damage growth model for a 1D body with:

$$\underline{C} = \underline{k} = \underline{1}, \quad \delta = 1, \quad s_0 = 0.02.$$

The different stress-strain curves are for different values of s_1 , indicating the transition from stable to unstable damage growth as s_1 is increased. The stress is computed using Eq. 58, to be discussed below.

In 2D, similar material and damage models (for a material with a bulk modulus of 10MPa) can simulate the stable accumulation of diffuse damage near a stress singularity, as shown in Fig. 3. The specimen contains a semicircular notch and is under combined normal and transverse loading corresponding to strain rates of $\dot{\epsilon}_{22} = 2.0\text{s}^{-1}$ and $\dot{\epsilon}_{12} = 1.0\text{s}^{-1}$. Until a strain of about $\epsilon_{22} = 0.12$, there is a stable growth of damage near the notch. Then there is a sudden transition to dynamic fracture. The crack rapidly propagates to the opposite free edge of the specimen. This transition from stable to unstable can be seen in the stress-strain curve shown in Fig. 4. This curve represents the total normal load in the vertical direction divided by the cross-sectional area of the specimen (on a cross-sectional plane that does not include the notch).

These examples demonstrate the potential usefulness of peridynamic damage mechanics in modeling materials that either fracture immediately or after a period of accumulated continuous damage. The compatibility of the peridynamic field equations with both continuous and discontinuous deformations is helpful in modeling the spontaneous nucleation and growth of fractures within a damaged material.

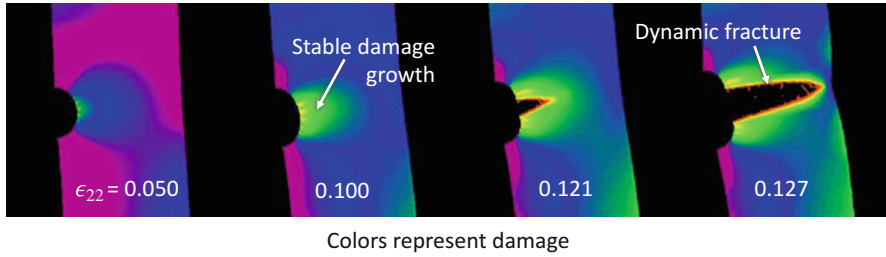


Fig. 3 Transition of stable damage growth at a stress concentration to dynamic fracture

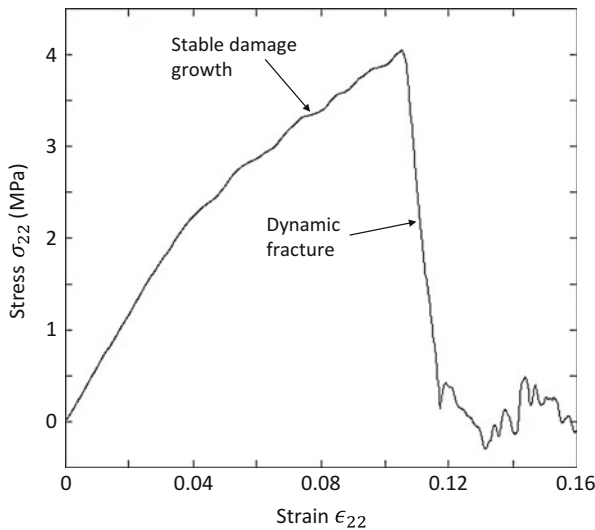


Fig. 4 Remote normal stress as a function of global strain for damage growth and fracture near a semicircular notch

Bond Breakage

A simpler approach to damage modeling in peridynamics has $\phi(\xi)$ jump discontinuously from 0 to 1 according to some criterion, which can be quite general. This approach, which is called *bond breakage*, is used in the vast majority of peridynamic codes because of its simplicity and reduced memory requirements. It has the disadvantage compared with continuously varying ϕ that instantaneous bond breakage can excite unwanted oscillations in a numerical grid. However, these can be suppressed in practice by applying damping forces to nodes after their bonds break.

A bond-breakage criterion can be as simple as a critical value of bond strain (Silling and Askari 2005). In this case, the critical bond strain for bond breakage can be calibrated to match a given critical energy release rate for the material. This bond-

breakage strain depends on the horizon as well as the critical energy release rate. Critical bond strain as a failure criterion is a natural way to model brittle fracture, particularly in mode I. For other modes, it becomes trickier to specify the critical strain, which may depend on the conditions in other bonds in the family. Nucleation of cracks is predicted by bond breakage damage growth laws, but with a simple bond-strain criterion, it is difficult to prescribe both the critical energy release rate and the critical conditions for crack nucleation simultaneously. Postfailure response of the bonds depends on the material model; for example, it may or may not be appropriate in a given material to allow bonds to sustain compressive force after they break.

In spite of the many successes of the bond breakage approach to peridynamic damage modeling, the above considerations help motivate the development of the more general approach using the thermodynamic force and failure surfaces as described in section “[Damage Evolution](#).”

Connections with the Local Theory

This section summarizes the mathematical and conceptual connections between the nonlocal peridynamic theory and the local theory. The connections discussed here include the relations between the force state and the stress tensor, between the deformation state and the deformation gradient tensor, between nonlocal heat transport and the Clausius-Duhem inequality, the scaling and convergence of material models, and local damage mechanics. Not discussed below but also important is the peridynamic version of the Eshelby-Rice J -integral (Hu et al. [2012b](#); Silling and Lehoucq [2010](#)).

Local Kinematics and Kinetics

Suppose a First Piola stress tensor σ is given. Let $\underline{\omega}$ be a positive scalar-valued state called the *influence function*. Consider the force state defined by:

$$\bar{\mathbf{T}}(\sigma) = \sigma \underline{\omega} \mathbf{K}^{-1} \underline{\mathbf{X}} \quad \text{or} \quad \bar{\mathbf{T}}(\sigma) \langle \xi \rangle = \sigma \underline{\omega} \langle \xi \rangle \mathbf{K}^{-1} \xi \quad \forall \xi \in \mathcal{H}_x \quad (57)$$

where \mathbf{K} is the *shape tensor* defined by:

$$\mathbf{K} = \int_{\mathcal{H}_x} \underline{\omega} \langle \xi \rangle \xi \otimes \xi \, dV_{\xi}.$$

The force state $\bar{\mathbf{T}}(\sigma)$ defined by Eq. [57](#) has the property that in a uniform deformation of a homogeneous body, the force per unit area across any plane transferred by all the bonds that cross this plane is equal to:

$$\boldsymbol{\tau} = \sigma \mathbf{n}$$

where \mathbf{n} is a unit normal to the plane. Conversely, for a given force state $\underline{\mathbf{T}}$, the stress tensor defined by:

$$\bar{\sigma}(\underline{\mathbf{T}}) = \int_{\mathcal{H}_x} \underline{\mathbf{T}}(\underline{\xi}) \otimes \underline{\xi} dV_{\xi} \tag{58}$$

has the same property. $\bar{\sigma}$ is called the *partial stress tensor*. For any σ ,

$$\bar{\sigma}(\bar{\underline{\mathbf{T}}}(\sigma)) = \sigma.$$

Analogous expressions can be derived for deformation states and deformation gradient tensors:

$$\bar{\underline{\mathbf{Y}}}(\mathbf{F}) \langle \underline{\xi} \rangle = \mathbf{F} \underline{\xi}, \quad \bar{\mathbf{F}}(\underline{\mathbf{Y}}) = \left(\int_{\mathcal{H}_x} \omega \langle \underline{\xi} \rangle \underline{\mathbf{Y}} \langle \underline{\xi} \rangle \otimes \underline{\xi} dV_{\xi} \right) \mathbf{K}^{-1} \quad \forall \mathbf{F}, \underline{\mathbf{Y}}. \tag{59}$$

It is easily confirmed that for any tensor \mathbf{F} ,

$$\bar{\mathbf{F}}(\bar{\underline{\mathbf{Y}}}(\mathbf{F})) = \mathbf{F}.$$

The relations (57) and (59) provide a way to adapt a local material model $\sigma(\mathbf{F})$ to peridynamics by setting

$$\underline{\mathbf{T}}(\underline{\mathbf{Y}}) = \bar{\underline{\mathbf{T}}}(\sigma(\bar{\mathbf{F}}(\underline{\mathbf{Y}}))) \tag{60}$$

for any $\underline{\mathbf{Y}}$. A peridynamic material model of the form (60) is called a *correspondence* model. These models have the properties that they are elastic, isotropic, and objective whenever the underlying local material model $\sigma(\mathbf{F})$ has these properties, in the sense of the local theory. Correspondence models generally exhibit zero-energy modes of deformation (Tupek and Radovitzky 2014) due to the noninvertibility of $\bar{\mathbf{F}}$, that is,

$$\bar{\mathbf{F}}(\underline{\mathbf{Y}}) = \bar{\mathbf{F}}(\underline{\mathbf{Y}}') \implies \underline{\mathbf{Y}} = \underline{\mathbf{Y}}'.$$

Several practical ways of reducing this type of instability in numerical models have been proposed. One such method penalizes the departures of the deformation state from a uniform deformation within the family. In this method, the material model is modified by including an additional term as follows:

$$\underline{\mathbf{T}}(\underline{\mathbf{Y}}) = \bar{\underline{\mathbf{T}}}(\sigma(\bar{\mathbf{F}}(\underline{\mathbf{Y}}))) + A \left\| \underline{\mathbf{Y}} - \bar{\mathbf{F}}(\underline{\mathbf{Y}}) \underline{\mathbf{X}} \right\|_{\underline{\mathbf{Y}}}$$

where the subscript denotes the Fréchet derivative and A is a constant (Silling 2017).

For any deformation (not necessarily uniform) of any peridynamic body (not necessarily homogeneous), the *peridynamic stress tensor* is defined by:

$$\nu(\mathbf{x}) = \int_{\mathcal{U}} \int_0^\infty \int_0^\infty (y+z)^2 (\underline{\mathbf{T}}[\mathbf{x}-z\mathbf{m}] \langle (y+z)\mathbf{m} \rangle) \otimes \mathbf{m} \, dy \, dz \, d\Omega_{\mathbf{m}}$$

where \mathcal{U} is the unit sphere and $d\Omega_{\mathbf{m}}$ is a differential spherical angle in the direction of the unit vector \mathbf{m} . ν has the surprising property (Lehoucq and Silling 2008; Lehoucq and von Lilienfeld 2010; Noll 1955) that

$$\nabla \cdot \nu = \int_{\mathcal{H}_x} \{ \underline{\mathbf{T}}[\mathbf{x}] \langle \mathbf{q} - \mathbf{x} \rangle - \underline{\mathbf{T}}[\mathbf{q}] \langle \mathbf{x} - \mathbf{q} \rangle \} dV_{\mathbf{q}}.$$

This means that the peridynamic equilibrium equation can be written as:

$$\nabla \cdot \nu + \mathbf{b} = 0,$$

which is formally the same as in the local theory.

The partial stress tensor $\bar{\sigma}$ is equal to ν in the special case of a uniform deformation of a homogeneous body. Otherwise, the sense in which it approximates ν is discussed in Silling et al. (2015).

Convergence of Peridynamics to the Local Theory

It seems reasonable to require that a proper nonlocal theory should converge, in some sense, to the local theory in the limit of “zero nonlocality.” To investigate how this convergence works in the peridynamic equations, we first consider how material models scale as the horizon decreases.

As the horizon is changed, holding the bulk properties of the material fixed, peridynamic material models satisfy certain scaling relations. Let $\underline{\mathbf{T}}_\delta$ and $\underline{\mathbf{Y}}_\delta$ denote the force state and the deformation state, respectively, for any $\delta > 0$. Let δ_1 and δ_2 be two values of the horizon and \mathcal{H}_1 and \mathcal{H}_2 the corresponding families. Suppose a material model $\widehat{\underline{\mathbf{T}}}_1$ is given for the horizon δ_1 . For any $\underline{\mathbf{Y}}_2$ on \mathcal{H}_2 , define the state $\underline{\mathbf{Y}}_1$ by:

$$\underline{\mathbf{Y}}_1 \langle \xi_1 \rangle = \frac{\delta_1}{\delta_2} \underline{\mathbf{Y}}_2 \left\langle \frac{\delta_2}{\delta_1} \xi_1 \right\rangle \quad \forall \xi_1 \in \mathcal{H}_1.$$

Consider the material model defined by:

$$\widehat{\underline{\mathbf{T}}}_2(\underline{\mathbf{Y}}_2) \langle \xi_2 \rangle = \left(\frac{\delta_1}{\delta_2} \right)^4 \widehat{\underline{\mathbf{T}}}_1(\underline{\mathbf{Y}}_1) \left\langle \frac{\delta_1}{\delta_2} \xi_2 \right\rangle \quad \forall \xi_2 \in \mathcal{H}_2. \tag{61}$$

It is easily confirmed that for a uniform deformation of a homogeneous body, $\bar{\sigma}$ defined in Eq. 58 is invariant with respect to this change in horizon:

$$\bar{\sigma} \left(\widehat{\mathbf{T}}_1 (\mathbf{Y}_1) \right) = \bar{\sigma} \left(\widehat{\mathbf{T}}_2 (\mathbf{Y}_2) \right).$$

In general, it can be shown that if a deformation is twice continuously differentiable, and if $\widehat{\mathbf{T}}_\delta$ scales according to Eq. 61, then the limit:

$$\sigma_0 := \lim_{\delta \rightarrow 0} \bar{\sigma} \left(\widehat{\mathbf{T}}_\delta (\mathbf{Y}_\delta) \right)$$

exists and that:

$$\lim_{\delta \rightarrow 0} \int_{\mathcal{H}_\delta} \left\{ \underline{\mathbf{T}}_\delta [\mathbf{x}] \langle \mathbf{q} - \mathbf{x} \rangle - \underline{\mathbf{T}}_\delta [\mathbf{q}] \langle \mathbf{x} - \mathbf{q} \rangle \right\} dV_{\mathbf{q}} = \nabla \cdot \sigma_0. \tag{62}$$

In summary, we now have a stress tensor field σ_0 such that, in the limit of zero horizon, the peridynamic accelerations equal the accelerations in the local theory computed from the divergence of σ_0 ; see Silling and Lehoucq (2008) for details. More rigorous results concerning convergence of peridynamics to local elasticity have been established (Emmrich et al. 2007).

Remarkably, Lipton has extended these results to discontinuous deformations; the limiting case of a peridynamic body containing a growing crack approaches a smooth solution in the local theory augmented by a Griffith crack that consumes energy at a definite rate as it grows. This result requires a peridynamic material model with a nonconvex strain-energy density function (Lipton 2014, 2016).

For heat transport, the statement analogous to Eq. 62 is as follows:

$$\lim_{\delta \rightarrow 0} \int_{\mathcal{H}_\delta} \left\{ \underline{H}_\delta [\mathbf{x}] \langle \mathbf{q} - \mathbf{x} \rangle - \underline{H}_\delta [\mathbf{q}] \langle \mathbf{x} - \mathbf{q} \rangle \right\} dV_{\mathbf{q}} = \nabla \cdot \mathbf{Q}_0 \tag{63}$$

where \mathbf{Q}_0 is the limiting heat flux vector field given by:

$$\mathbf{Q}_0 := \lim_{\delta \rightarrow 0} \bar{\mathbf{Q}}_\delta, \quad \bar{\mathbf{Q}}_\delta := \bar{\mathbf{Q}} (\underline{H}_\delta) = - \int_{\mathcal{H}_x} \underline{H}_\delta \langle \xi \rangle \xi dV_\xi. \tag{64}$$

The minus sign appears by convention in the second of Eq. 64 so that the heat flux $\bar{\mathbf{Q}} \cdot \mathbf{n}$ through a plane normal to a unit vector \mathbf{n} will be positive if energy is flowing parallel to \mathbf{n} , rather than opposite to it. Recall the inequality (23) derived from the second law,

$$\underline{\beta} \bullet \underline{H}_\delta \leq 0, \quad \underline{\beta} [\mathbf{x}] \langle \mathbf{q} - \mathbf{x} \rangle = \frac{1}{\theta(\mathbf{q})} - \frac{1}{\theta(\mathbf{x})}. \tag{65}$$

For a smooth temperature field, as $\delta \rightarrow 0$, we can use the first term of a Taylor series to write down the first-order approximation:

$$\underline{\beta} \langle \underline{\xi} \rangle \approx \underline{\xi} \cdot \nabla \beta(\mathbf{x}) \quad \forall \underline{\xi} \in \mathcal{H}_x,$$

hence Eq. 65 can be approximated by:

$$0 \geq \underline{\beta} \bullet \underline{H}_\delta \approx \int_{\mathcal{H}_x} \underline{\xi} \cdot (\nabla \beta) \underline{H}_\delta \langle \underline{\xi} \rangle = -\nabla \beta \cdot \overline{\mathbf{Q}}_\delta. \tag{66}$$

Similarly, for small δ , Eq. 63 is approximated by:

$$h_\delta := \int_{\mathcal{H}_\delta} \{ \underline{H}_\delta[\mathbf{x}] \langle \mathbf{q} - \mathbf{x} \rangle - \underline{H}_\delta[\mathbf{q}] \langle \mathbf{x} - \mathbf{q} \rangle \} dV_{\mathbf{q}} \approx \int_{\mathcal{H}_x} \underline{\xi} \cdot \nabla \underline{H}_\delta \langle \underline{\xi} \rangle = -\nabla \cdot \overline{\mathbf{Q}}_\delta. \tag{67}$$

From Eqs. 20, 66, and 67,

$$\begin{aligned} \dot{\eta} &\geq \beta h_\delta + \beta s \\ &\approx -\beta \nabla \cdot \overline{\mathbf{Q}}_\delta + \beta s \\ &\geq -\beta \nabla \cdot \overline{\mathbf{Q}}_\delta - \nabla \beta \cdot \overline{\mathbf{Q}}_\delta + \beta s \\ &= -\nabla \cdot (\beta \overline{\mathbf{Q}}_\delta) + \beta s. \end{aligned}$$

Omitting some of the details of taking the limit, the result for $\delta \rightarrow 0$ is:

$$\dot{\eta} \geq \frac{s}{\theta} - \nabla \cdot \left(\frac{\mathbf{Q}_0}{\theta} \right).$$

This is a form of the Clausius-Duhem inequality of local continuum thermodynamics. It is interesting that this local form of the second law with heat transport can be derived from the peridynamic version of the second law without specifying any particular form of the constitutive model $\widehat{H}(\Theta)$ or assuming a particular physical mechanism for heat transport (conduction, convection, radiation, etc.).

Local Continuum Damage Mechanics

The failure characteristics of engineering materials are frequently expressed in terms of failure surfaces, with or without some specification of postfailure behavior. These expressions can sometimes be converted to peridynamic failure surfaces using the partial stress and approximate deformation gradient defined in Eq. 58 and 59 above. For example, if a failure surface in terms of the stress tensor and deformation

gradient tensor is given in the form $\mathcal{S}^0(\boldsymbol{\sigma}, \mathbf{F}, \varphi)$, where φ is a scalar damage variable, we can define a peridynamic failure surface by:

$$\mathcal{S}(\underline{\mathbf{T}}, \underline{\mathbf{Y}}, \underline{\phi}) = \mathcal{S}^0(\bar{\boldsymbol{\sigma}}(\underline{\mathbf{T}}), \bar{\mathbf{F}}(\underline{\mathbf{Y}}), \bar{\varphi}(\underline{\phi})), \quad \bar{\varphi}(\underline{\phi}) = \frac{\underline{\omega} \bullet \underline{\phi}}{\underline{\omega} \bullet \underline{\mathbf{1}}}.$$

Evaluating the required Fréchet derivatives in Eq. 54 then leads to:

$$\mathcal{S}_{\underline{\mathbf{T}}} = \frac{\partial \mathcal{S}^0}{\partial \boldsymbol{\sigma}} \underline{\mathbf{X}}, \quad \mathcal{S}_{\underline{\mathbf{Y}}} = \frac{\partial \mathcal{S}^0}{\partial \mathbf{F}} \underline{\omega}(\underline{\xi}) \mathbf{K}^{-1} \underline{\mathbf{X}}, \quad \mathcal{S}_{\underline{\phi}} = \frac{\partial \mathcal{S}^0}{\partial \varphi} \frac{\underline{\omega}}{\underline{\omega} \bullet \underline{\mathbf{1}}}.$$

The remaining derivatives that appear in Eq. 54 are obtained from the peridynamic material model, as before.

Discussion

Why and how does a crack grow? How does a continuous body become discontinuous? Why does nature seem to favor these discontinuities as energy minimizers, yet equip real materials with energy barriers that resist their formation and growth? How can real materials be designed or optimized to resist cracking? These and other fundamental questions possibly can be studied within the peridynamic theory.

Although peridynamics is often used to model brittle fracture and fragmentation (for example, (Hu et al. 2013)), diverse new applications are continually being discovered. These have recently included, for example:

- Electromigration in integrated circuits (Gerstle et al. 2008; Oterkus et al. 2013)
 - Biological cell mechanics and tumor growth (Lejeune and Linder 2017a, b; Taylor et al. 2016)
 - Damage in materials due to high voltage breakdown (Wildman and Gazonas 2015)
 - Effects of residual thermal stress on fracture in glass (Jeon et al. 2015; Kilic and Madenci 2009)
 - Failure of reinforced concrete (Gerstle et al. 2010)
 - Mechanics of nanocomposites (Prakash and Seidel 2016)
 - Fluid transport and hydraulic fracture in rocks (Katiyar et al. 2014; Nadimi 2015; Ouchi et al. 2015; Van Der Merwe 2014)
 - Solitons (Silling 2016)
 - Corrosion (Chen and Bobaru 2015)
- and many others.

Acknowledgment Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the US Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

References

- F. Bobaru, M. Duangpanya, The peridynamic formulation for transient heat conduction. *Int. J. Heat Mass Transf.* **53**, 4047–4059 (2010)
- F. Bobaru, M. Duangpanya, A peridynamic formulation for transient heat conduction in bodies with evolving discontinuities. *J. Comput. Phys.* **231**, 2764–2785 (2012)
- Z. Chen, F. Bobaru, Peridynamic modeling of pitting corrosion damage. *J. Mech. Phys. Solids* **78**, 352–381 (2015)
- S.R. Chowdhury, P. Roy, D. Roy, J. Reddy, A peridynamic theory for linear elastic shells. *Int. J. Solids Struct.* **84**, 110–132 (2016)
- B.D. Coleman, W. Noll, The thermodynamics of elastic materials with heat conduction and viscosity. *Arch. Ration. Mech. Anal.* **13**, 167–178 (1963)
- C. Diyaroglu, E. Oterkus, S. Oterkus, E. Madenci, Peridynamics for bending of beams and plates with transverse shear deformation. *Int. J. Solids Struct.* **69**, 152–168 (2015)
- E. Emmrich, O. Weckner, et al., On the well-posedness of the linear peridynamic model and its convergence towards the navier equation of linear elasticity. *Commun. Math. Sci.* **5**, 851–864 (2007)
- J.T. Foster, S.A. Silling, W.W. Chen, Viscoplasticity using peridynamics. *Int. J. Numer. Methods Eng.* **81**, 1242–1258 (2010)
- E. Fried, New insights into the classical mechanics of particle systems. *Discrete Contin. Dyn. Syst.* **28**, 1469–1504 (2010)
- W. Gerstle, N. Sau, S.A. Silling, Peridynamic modeling of concrete structures. *Nucl. Eng. Des.* **237**, 1250–1258 (2007)
- W. Gerstle, S. Silling, D. Read, V. Tewary, R. Lehoucq, Peridynamic simulation of electromigration. *Comput. Mater. Continua* **8**, 75–92 (2008)
- W. Gerstle, N. Sakhavand, S. Chapman, Peridynamic and continuum models of reinforced concrete lap splice compared, in *Fracture Mechanics of Concrete and Concrete Structures, Recent Advances in Fracture Mechanics of Concrete*, ed. by B.H. Oh, et al. (2010), pp. 306–312
- J. O’Grady, J. Foster, Peridynamic plates and flat shells: a non-ordinary, state-based model. *Int. J. Solids Struct.* **51**, 4572–4579 (2014)
- M.E. Gurtin, W.O. Williams, On the first law of thermodynamics. *Arch. Ration. Mech. Anal.* **42**, 77–92 (1971)
- M.E. Gurtin, E. Fried, L. Anand, *The mechanics and thermodynamics of continua* (Cambridge University Press, Cambridge, 2010), pp. 232–233
- W. Hu, Y.D. Ha, F. Bobaru, Peridynamic model for dynamic fracture in unidirectional fiber-reinforced composites. *Comput. Methods Appl. Mech. Eng.* **217**, 247–261 (2012a)
- W. Hu, Y.D. Ha, F. Bobaru, S.A. Silling, The formulation and computation of the nonlocal J-integral in bond-based peridynamics. *Int. J. Fract.* **176**, 195–206 (2012b)
- W. Hu, Y. Wang, J. Yu, C.-F. Yen, F. Bobaru, Impact damage on a thin glass plate with a thin polycarbonate backing. *Int. J. Impact Eng.* **62**, 152–165 (2013)
- B. Jeon, R.J. Stewart, I.Z. Ahmed, Peridynamic simulations of brittle structures with thermal residual deformation: strengthening and structural reactivity of glasses under impacts. *Proc. R. Soc. A* **471**, 20150231. (2015)
- A. Katiyar, J.T. Foster, H. Ouchi, M.M. Sharma, A peridynamic formulation of pressure driven convective fluid transport in porous media. *J. Comput. Phys.* **261**, 209–229 (2014)
- B. Kilic, E. Madenci, Prediction of crack paths in a quenched glass plate by using peridynamic theory. *Int. J. Fract.* **156**, 165–177 (2009)
- R.B. Lehoucq, M.P. Sears, Statistical mechanical foundation of the peridynamic nonlocal continuum theory: energy and momentum conservation laws. *Phys. Rev. E* **84**, 031112 (2011)
- R.B. Lehoucq, S.A. Silling, Force flux and the peridynamic stress tensor. *J. Mech. Phys. Solids* **56**, 1566–1577 (2008)
- R.B. Lehoucq, O.A. von Lilienfeld, Translation of Walter Noll’s derivation of the fundamental equations of continuum thermodynamics from statistical mechanics. *J. Elast.* **100**, 5–24 (2010)

- E. Lejeune, C. Linder, Modeling tumor growth with peridynamics. *Biomech. Model. Mechanobiol.*, 1–17 (2017a)
- E. Lejeune, C. Linder, Quantifying the relationship between cell division angle and morphogenesis through computational modeling. *J. Theor. Biol.* **418**, 1–7 (2017b)
- R. Lipton, Dynamic brittle fracture as a small horizon limit of peridynamics. *J. Elast.* **117**, 21–50 (2014)
- R. Lipton, Cohesive dynamics and brittle fracture. *J. Elast.* **142**, 1–49 (2016)
- E. Madenci, S. Oterkus, Ordinary state-based peridynamics for plastic deformation according to von Mises yield criteria with isotropic hardening. *J. Mech. Phys. Solids* **86**, 192–219 (2016)
- J.A. Mitchell, A non-local, ordinary-state-based viscoelasticity model for peridynamics. Technical report SAND2011-8064, Sandia National Laboratories, Albuquerque/Livermore, October 2011a
- J.A. Mitchell, A nonlocal, ordinary, state-based plasticity model for peridynamics. Technical report SAND2011-3166, Sandia National Laboratories, Albuquerque/Livermore, October 2011b
- S. Nadimi, State-based peridynamics simulation of hydraulic fracture phenomenon in geological media. Master's thesis, The University of Utah, 2015
- W. Noll, Die Herleitung der Grundgleichungen der Thermomechanik der Kontinua aus der statistischen Mechanik. *J. Ration. Mech. Anal.* **4**, 627–646 (1955.) In German, English translation available
- E. Oterkus, E. Madenci, Peridynamic analysis of fiber-reinforced composite materials. *J. Mech. Mater. Struct.* **7**, 45–84 (2012)
- S. Oterkus, J. Fox, E. Madenci, Simulation of electro-migration through peridynamics, in *2013 IEEE 63rd Electronic Components and Technology Conference* (IEEE, 2013), pp. 1488–1493
- S. Oterkus, E. Madenci, A. Agwai, Fully coupled peridynamic thermomechanics. *J. Mech. Phys. Solids* **64**, 1–23 (2014a)
- S. Oterkus, E. Madenci, A. Agwai, Peridynamic thermal diffusion. *J. Comput. Phys.* **265**, 71–96 (2014b)
- H. Ouchi, A. Katiyar, J. Foster, M.M. Sharma, et al., A peridynamics model for the propagation of hydraulic fractures in heterogeneous, naturally fractured reservoirs. in *SPE Hydraulic Fracturing Technology Conference* (Society of Petroleum Engineers, 2015)
- N. Prakash, G.D. Seidel, A coupled electromechanical peridynamics framework for modeling carbon nanotube reinforced polymer composites, in *57th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, p. 0936, (2016)
- S.A. Silling, Reformulation of elasticity theory for discontinuities and long-range forces. *J. Mech. Phys. Solids* **48**, 175–209 (2000)
- S.A. Silling, Linearized theory of peridynamic states. *J. Elast.* **99**, 85–111 (2010)
- S.A. Silling, Solitary waves in a peridynamic elastic solid. *J. Mech. Phys. Solids* **96**, 121–132 (2016)
- S.A. Silling, Stability of peridynamic correspondence material models and their particle discretizations. *Comput. Methods Appl. Mech. Eng.* **322**, 42–57 (2017)
- S.A. Silling, E. Askari, A meshfree method based on the peridynamic model of solid mechanics. *Comput. Struct.* **83**, 1526–1535 (2005)
- S.A. Silling, R.B. Lehoucq, Convergence of peridynamics to classical elasticity theory. *J. Elast.* **93**, 13–37 (2008)
- S.A. Silling, R.B. Lehoucq, The peridynamic theory of solid mechanics. *Adv. Appl. Mech.* **44**, 73–166 (2010)
- S.A. Silling, M. Epton, O. Weckner, J. Xu, E. Askari, Peridynamic states and constitutive modeling. *J. Elast.* **88**, 151–184 (2007)
- S.A. Silling, D. Littlewood, P. Seleson, Variable horizon in a peridynamic medium. *J. Mech. Mater. Struct.* **10**, 591–612 (2015)
- S.A. Silling, M.L. Parks, J.R. Kamm, O. Weckner, M. Rassaian, Modeling shockwaves and impact phenomena with Eulerian peridynamics. *Int. J. Impact Eng.* **107**, 47–57 (2017)
- S. Sun, V. Sundararaghavan, A peridynamic implementation of crystal plasticity. *Int. J. Solids Struct.* **51**, 3350–3360 (2014)

- M. Taylor, I. Gözen, S. Patel, A. Jesorka, K. Bertoldi, Peridynamic modeling of ruptures in biomembranes. *PLoS One* **11**, e0165947 (2016)
- M. Tupek, R. Radovitzky, An extended constitutive correspondence formulation of peridynamics based on nonlinear bond-strain measures. *J. Mech. Phys. Solids* **65**, 82–92 (2014)
- C.W. Van Der Merwe, A peridynamic model for sleeved hydraulic fracture. Master's thesis, Stellenbosch University, Stellenbosch, (2014)
- T.L. Warren, S.A. Silling, A. Askari, O. Weckner, M.A. Epton, J. Xu, A nonordinary state-based peridynamic method to model solid material deformation and fracture. *Int. J. Solids Struct.* **46**, 1186–1195 (2009)
- O. Weckner, N.A.N. Mohamed, Viscoelastic material models in peridynamics. *Appl. Math. Comput.* **219**, 6039–6043 (2013)
- R. Wildman, G. Gazonas, A dynamic electro-thermo-mechanical model of dielectric breakdown in solids using peridynamics. *J. Mech. Mater. Struct.* **10**, 613–630 (2015)



Recent Progress in Mathematical and Computational Aspects of Peridynamics

35

Marta D'Elia, Qiang Du, and Max Gunzburger

Contents

Introduction	1198
Peridynamic Models and Their Linearized Form	1199
Mathematical Framework	1200
A Nonlocal Vector Calculus	1200
Formulating Nonlocal Models via Nonlocal Operators	1201
A Linear Steady-State Problem as an Illustration	1202
Variational Formulation	1204
Discretization Schemes	1205
Peridynamics as a Multiscale Mono-model	1208
A Multiscale Finite Element Implementation of Peridynamics	1210
Control, Identification, and Obstacle Problems	1212
Nonlocal Optimal Control Problems	1212
Nonlocal Parameter Identification Problems	1214
Nonlocal Obstacle Problems	1217
References	1219

M. D'Elia (✉)

Optimization and Uncertainty Quantification Department Center for Computing Research, Sandia National Laboratories, Albuquerque, NM, USA

e-mail: mdelia@sandia.gov

Q. Du

Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY, USA

e-mail: qd2125@columbia.edu

M. Gunzburger

Department of Scientific Computing, Florida State University, Tallahassee, FL, USA

e-mail: mgunzburger@fsu.edu

© This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply 2019

G. Z. Voyiadjis (ed.), *Handbook of Nonlocal Continuum Mechanics for Materials and Structures*, https://doi.org/10.1007/978-3-319-58729-5_30

1197

Abstract

Recent developments in the mathematical and computational aspects of the nonlocal peridynamic model for material mechanics are provided. Based on a recently developed vector calculus for nonlocal operators, a mathematical framework is constructed that has proved useful for the mathematical analyses of peridynamic models and for the development of finite element discretizations of those models. A specific class of discretization algorithms referred to as asymptotically compatible schemes is discussed; this class consists of methods that converge to the proper limits as grid sizes and nonlocal effects tend to zero. Then, the multiscale nature of peridynamics is discussed including how, as a single model, it can account for phenomena occurring over a wide range of scales. The use of this feature of the model is shown to result in efficient finite element implementations. In addition, the mathematical and computational frameworks developed for peridynamic simulation problems are shown to extend to control, coefficient identification, and obstacle problems.

Keywords

Peridynamics · Nonlocal vector calculus · Variational forms · Multiscale methods · Finite element method · Optimal control · Obstacle problems

Introduction

The peridynamic (PD) model for solid mechanics was introduced in Silling (2000), followed by a more generally applicable version in Silling et al. (2007); see also the review (Silling and Lehoucq 2010). The main features of the model are that it is *a continuum model that is free from spatial derivatives, allows for nonlocal interactions, and allows for solutions that contain jump discontinuities across lower-dimensional manifolds*. As such, it is especially well suited for simulations of material failure phenomena such as fracture. Indeed, despite its relatively recent development, the effectiveness of PD has already been demonstrated in several sophisticated applications, including the fracture and failure of composites, crack instability, fracture of polycrystals, and nanofiber networks; see Askari et al. (2008) for a review and also Bobaru and Silling (2004), Bobaru et al. (2005), Gerstle and Sau (2004), Gerstle et al. (2005), Littlewood (2010), Silling (2003), Silling and Askari (2004), Silling and Bobaru (2005), Silling et al. (2003), Weckner and Abeyaratne (2005), and Weckner and Emmrich (2005). The successful application of PD for multiscale engineering analyses has been enabled by the development of software packages, see, e.g., Parks et al. (2008, 2012) and Silling and Askari (2005). There has also been extensive studies of peridynamic models appearing in the engineering, material science, and mathematical literatures.

An additional feature of PD is that, all on its own, it is a *multiscale* model. A *valid* material model is one that provides a faithful description of the physical phenomena. A *tractable* material model is one for which useful information can be extracted,

e.g., through discretization, in an efficient manner. These two requirements allow us to define a *multiscale* material model as one that is valid and tractable over a wide range of spatial and temporal scales. In most cases, multiscale models are defined by coupling two different models that operate at different scales, e.g., molecular dynamics and classical elasticity; peridynamics provides the opportunity to do away with the need for such troublesome couplings. Thus, we refer to peridynamics as a multiscale *mono*-model in contrast to multiscale *multi*-models for which two different models, e.g., classical elasticity and molecular dynamics, are coupled to produce a multiscale model. This feature, which is discussed below, results from the introduction of a length scale $\varepsilon > 0$, often referred to as the horizon, such that points separated by a distance greater than ε do not, for all practical purposes, interact.

The rest of this section is devoted to a very brief description of peridynamic models. Then, in section “[Mathematical Framework](#)” we present a mathematical framework, based on a recently developed vector calculus for nonlocal operators, that has proved useful for the mathematical analyses of peridynamic models and for the development of finite element discretizations of those models. In particular, the recently developed concept of asymptotically compatible schemes is discussed. In section “[Peridynamics as a Multiscale Mono-model](#),” the multiscale nature of peridynamics as a mono-model for material mechanics is discussed as is the implications that feature of peridynamics has on efficient finite element implementations. In section “[Control, Identification, and Obstacle Problems](#),” we show how the mathematical and computational frameworks developed for peridynamic simulation problems can be extended to control, coefficient identification, and obstacle problems. The material presented in this paper is drawn from, among other sources, Du et al. (2012a, 2013a,b) and Du (2016a,b) for section “[Mathematical Framework](#),” Xu et al. (2016b) for section “[Peridynamics as a Multiscale Mono-model](#),” and D’Elia and Gunzburger (2014, 2016) and Guan and Gunzburger (2017) for section “[Control, Identification, and Obstacle Problems](#).”

Peridynamic Models and Their Linearized Form

The general state-based peridynamic equation of motion (Silling et al. 2007) takes the form

$$\rho \partial_{tt} \mathbf{u}(\mathbf{x}, t) = \int_{\mathbb{R}^n} \{ \underline{\mathbb{T}}[\mathbf{x}] \langle \mathbf{y} - \mathbf{x} \rangle - \underline{\mathbb{T}}[\mathbf{y}] \langle \mathbf{x} - \mathbf{y} \rangle \} d\mathbf{y} + \mathbf{b}(\mathbf{x}) \quad (1)$$

with $\underline{\mathbb{T}}[\mathbf{x}] \langle \mathbf{y} - \mathbf{x} \rangle$ and $\underline{\mathbb{T}}[\mathbf{y}] \langle \mathbf{x} - \mathbf{y} \rangle$ denoting the peridynamic force states, \mathbf{u} the displacement vector, ρ the material density, and \mathbf{b} a given body force density. Under the assumption of small deformation, (1) can be approximated by the linear integrodifferential equation

$$\rho \partial_{tt} \mathbf{u}(\mathbf{x}, t) = \int_{\mathbb{R}^n} C_\varepsilon(\mathbf{y}, \mathbf{x}) (\mathbf{u}(\mathbf{y}, t) - \mathbf{u}(\mathbf{x}, t)) d\mathbf{y} + \mathbf{b}(\mathbf{x}). \quad (2)$$

As an example, we have the linear bond-based model for which C_ε is a rank-one tensor of the form

$$C_\varepsilon(\mathbf{y}, \mathbf{x}) = \omega_\varepsilon(|\mathbf{y} - \mathbf{x}|) [\boldsymbol{\alpha}(\mathbf{y}, \mathbf{x}) \otimes \boldsymbol{\alpha}(\mathbf{y}, \mathbf{x})], \quad \text{where} \quad \boldsymbol{\alpha}(\mathbf{y}, \mathbf{x}) = \frac{\mathbf{y} - \mathbf{x}}{|\mathbf{y} - \mathbf{x}|^2} \quad (3)$$

denotes a scaled vector along the bond between \mathbf{x} and \mathbf{y} . We assume that nonlocal interactions occur only when the distance between \mathbf{x} and \mathbf{y} is smaller than a specified distance ε , referred to as the *horizon*. Thus, the nonlocal interaction kernel $\omega_\varepsilon(|\mathbf{y} - \mathbf{x}|)L$ is supported in a spherical neighborhood of radius ε . However, for the state-based peridynamic model, the range of interactions becomes broader due to the indirect interactions (Silling 2010). Nevertheless, we still use ε as a characteristic measure of the range of nonlocal interactions.

Local limit of peridynamic models. In peridynamic models, if $\varepsilon \rightarrow 0$, then nonlocal interactions become localized. We expect that in several settings, such as linear peridynamic models, the nonlocal models reduce to classical partial differential equation (PDE) models under proper assumptions on the interaction kernels. Intuitively, these assumptions imply that $\omega_\varepsilon(|\mathbf{z}|)$ approaches a constant multiple of the Dirac delta function at $\mathbf{z} = \mathbf{0}$. For the linear bond-based peridynamics, such studies have been carried out in Du et al. (2013b), Du and Zhou (2011), and Mengesha and Du (2013).

Mathematical Framework

Understanding and developing mathematical treatments of nonlocal models call for a new mathematical framework. In this section, we give a brief description of that framework and its use in the nonlocal mechanic setting.

A Nonlocal Vector Calculus

Because the quantities of interest in continuum mechanics are mostly vector and tensor-valued function, the nonlocal vector calculus presented in Du et al. (2013a) provides a systematic means for formulating nonlocal mechanical models such as peridynamics. Nonlocal operators are defined as the basic building blocks of the nonlocal vector calculus. For example, for any two points \mathbf{y} and \mathbf{x} in \mathbb{R}^n , we choose a two-point vector $\boldsymbol{\alpha}(\mathbf{x}, \mathbf{y}): \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^m$ that is antisymmetric, i.e., we have $\boldsymbol{\alpha}(\mathbf{y}, \mathbf{x}) = -\boldsymbol{\alpha}(\mathbf{x}, \mathbf{y})$, and then, respectively, define a nonlocal divergence operator \mathcal{D} acting on tensors and its adjoint operator \mathcal{D}^* acting on vectors by

$$\begin{aligned} (\mathcal{D}\underline{\boldsymbol{\psi}})(\mathbf{x}) &= \int_{\mathbb{R}^n} (\underline{\boldsymbol{\psi}}(\mathbf{x}, \mathbf{y}) + \underline{\boldsymbol{\psi}}(\mathbf{y}, \mathbf{x}))\boldsymbol{\alpha}(\mathbf{x}, \mathbf{y})d\mathbf{y} & \forall \underline{\boldsymbol{\psi}}(\mathbf{x}, \mathbf{y}): \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m} \\ (\mathcal{D}^*\mathbf{v})(\mathbf{x}, \mathbf{y}) &= (\mathbf{v}(\mathbf{y}) - \mathbf{v}(\mathbf{x})) \otimes \boldsymbol{\alpha}(\mathbf{y}, \mathbf{x}) & \forall \mathbf{v}(\mathbf{y}): \mathbb{R}^n \rightarrow \mathbb{R}^m. \end{aligned}$$

\mathfrak{D} and \mathfrak{D}^* are adjoint operators in the sense that

$$\int_{\mathbb{R}^n} \mathbf{v}(\mathbf{x}) \cdot (\mathfrak{D}\underline{\psi})(\mathbf{x})d\mathbf{x} = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} (\mathfrak{D}^*\mathbf{v})(\mathbf{x}, \mathbf{y})\underline{\psi}(\mathbf{x}, \mathbf{y})d\mathbf{y}d\mathbf{x}.$$

Note that $\mathfrak{D}^*\mathbf{v}$ is a symmetric two-point tensor-valued function, i.e., $(\mathfrak{D}^*\mathbf{v})(\mathbf{x}, \mathbf{y}) = (\mathfrak{D}^*\mathbf{v})(\mathbf{y}, \mathbf{x})$. On the other hand, for any two-point tensor function $\underline{\psi}(\mathbf{x}, \mathbf{y})$, $\mathfrak{D}\underline{\psi}$ depends only on the symmetric part of $\underline{\psi}$ with respect to its two variables \mathbf{x} and \mathbf{y} . With respect to the dimension parameters m and n , the cases most relevant to peridynamics are $m = 1$ and $m = n$.

For notational convenience, as in Du et al. (2013a), we introduce the nonlocal divergence operator \mathcal{D} acting on vectors and its adjoint operator \mathcal{D}^* acting on scalars by

$$\begin{aligned} (\mathcal{D}\boldsymbol{\varphi})(\mathbf{x}) &= \int_{\mathbb{R}^n} (\boldsymbol{\varphi}(\mathbf{x}, \mathbf{y}) + \boldsymbol{\varphi}(\mathbf{y}, \mathbf{x})) \cdot \boldsymbol{\alpha}(\mathbf{x}, \mathbf{y})d\mathbf{y} & \forall \boldsymbol{\varphi}: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n \\ (\mathcal{D}^*v)(\mathbf{x}, \mathbf{y}) &= (v(\mathbf{y}) - v(\mathbf{x}))\boldsymbol{\alpha}(\mathbf{y}, \mathbf{x}) & \forall v: \mathbb{R}^n \rightarrow \mathbb{R}. \end{aligned}$$

To account for the indirect interactions between two material points present in the state-based peridynamic models, the weighted nonlocal operators $\mathfrak{D}_{\omega_\varepsilon}$ and $\mathfrak{D}_{\omega_\varepsilon}^*$ are introduced in Du et al. (2013a). For example, letting $\omega_\varepsilon(\mathbf{x}, \mathbf{y}): \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ denote a nonnegative scalar-valued two-point function and given the function $\mathbf{u}(\mathbf{x}): \mathbb{R}^n \rightarrow \mathbb{R}^m$, we have the weighted divergence operator given by

$$\mathfrak{D}_{\omega_\varepsilon}(\mathbf{u})(\mathbf{x}) = \mathfrak{D}(\omega_\varepsilon(\mathbf{x}, \mathbf{y})\mathbf{u}(\mathbf{x})) \quad \text{for } \mathbf{x} \in \mathbb{R}^n.$$

These operators are more closely aligned with conventional first-order differential operators defined on vector fields compared to the operators \mathcal{D} , \mathcal{D}^* , \mathfrak{D} , and \mathfrak{D}^* . More rigorous definitions and careful modifications in bounded domains are given in Mengesha and Du (2016). Additional formal discussions about \mathcal{D} , \mathcal{D}^* , $\mathfrak{D}_{\omega_\varepsilon}$, and $\mathfrak{D}_{\omega_\varepsilon}^*$ can be found in Du et al. (2013a) and rigorous derivations in Mengesha and Du (2016) along with nonlocal Green’s identities such as

$$\int_{\mathbb{R}^n} \mathbf{u}(\mathfrak{D}(\omega(\mathfrak{D}^*\mathbf{v})))d\mathbf{x} - \int_{\mathbb{R}^n} (\mathfrak{D}(\omega(\mathfrak{D}^*\mathbf{u})))\mathbf{v}d\mathbf{x} = 0.$$

The nonlocal Green’s identities are nonlocal analogs of the classical counterparts and are crucial in defining variational formulations of nonlocal mechanic problems.

Formulating Nonlocal Models via Nonlocal Operators

The nonlocal vector calculus allows us to reformulate linear peridynamic models in ways similar to the way local linear PDE models can be expressed in terms of basic divergence, gradient, and curl operators. For example, for a given kernel $\omega_\varepsilon =$

$\omega_\varepsilon(|\mathbf{y} - \mathbf{x}|)$ and any function $\mathbf{u} = \mathbf{u}(\mathbf{x})$ defined in \mathbb{R}^n , using the matrix and vector product identities

$$(\boldsymbol{\beta} \otimes \boldsymbol{\alpha})\boldsymbol{\alpha} = \boldsymbol{\beta}\boldsymbol{\alpha}^T\boldsymbol{\alpha} = |\boldsymbol{\alpha}|^2\boldsymbol{\beta} \quad \text{and} \quad (\boldsymbol{\alpha} \otimes \boldsymbol{\alpha})\boldsymbol{\beta} = \boldsymbol{\alpha}^T\boldsymbol{\beta}\boldsymbol{\alpha} \quad \forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{R}^m \quad (4)$$

and the notation $\underline{\omega}_\varepsilon(|\mathbf{y} - \mathbf{x}|) = \omega_\varepsilon(|\mathbf{y} - \mathbf{x}|)|\boldsymbol{\alpha}(\mathbf{x}, \mathbf{y})|^2$, a composition of the operators \mathcal{D} and \mathcal{D}^* yields

$$\mathcal{L}_\varepsilon(\mathbf{u})(\mathbf{x}) = -\mathcal{D}(\omega_\varepsilon \mathcal{D}^*\mathbf{u})(\mathbf{x}) = \int_{\mathbb{R}^n} \underline{\omega}_\varepsilon(|\mathbf{y} - \mathbf{x}|) ((\mathbf{u}(\mathbf{y}) - \mathbf{u}(\mathbf{x})) \otimes \boldsymbol{\alpha}(\mathbf{y}, \mathbf{x})) \boldsymbol{\alpha}(\mathbf{y}, \mathbf{x}) d\mathbf{y} \quad (5)$$

which corresponds to the linear Navier operator for bond-based peridynamics. On the other hand, for a scalar function u , we have

$$\mathcal{L}_\varepsilon(u)(\mathbf{x}) = -\mathcal{D}(\omega_\varepsilon \mathcal{D}^*u)(\mathbf{x}) = \int_{\mathbb{R}^n} \underline{\omega}_\varepsilon(|\mathbf{y} - \mathbf{x}|)(u(\mathbf{y}) - u(\mathbf{x})) d\mathbf{y} \quad (6)$$

which is often referred to a *nonlocal Laplacian* or a nonlocal diffusion operator.

A Linear Steady-State Problem as an Illustration

We consider a linear steady-state problem related to the nonlocal bond-based peridynamic model for which the nonlocal force is a linear Hookean spring force aligned with the bond direction. We have

$$-\mathcal{L}_\varepsilon\mathbf{u}(\mathbf{x}) = \mathbf{b}(\mathbf{x}) \quad \forall \mathbf{x} \in \Omega, \quad (7)$$

where Ω denotes a bounded domain in \mathbb{R}^n and the nonlocal bond-based peridynamic operator \mathcal{L}_ε is defined by

$$\begin{aligned} \mathcal{L}_\varepsilon\mathbf{u}(\mathbf{x}) &= -\mathcal{D}(\omega_\varepsilon \mathcal{D}^*(\mathbf{u}))(\mathbf{x}) \\ &= \int_{\Omega \cup \Omega_I} \omega_\varepsilon(|\mathbf{y} - \mathbf{x}|)\boldsymbol{\alpha}(\mathbf{y}, \mathbf{x}) \otimes \boldsymbol{\alpha}(\mathbf{y}, \mathbf{x})(\mathbf{u}(\mathbf{y}) - \mathbf{u}(\mathbf{x})) d\mathbf{y} \\ &= \int_{\Omega \cup \Omega_I} \omega_\varepsilon(|\mathbf{y} - \mathbf{x}|) \frac{\mathbf{y} - \mathbf{x}}{|\mathbf{y} - \mathbf{x}|^2} \otimes \frac{\mathbf{y} - \mathbf{x}}{|\mathbf{y} - \mathbf{x}|^2} (\mathbf{u}(\mathbf{y}) - \mathbf{u}(\mathbf{x})) d\mathbf{y}. \end{aligned} \quad (8)$$

Here, Ω_I denotes the *interaction domain* which is defined as containing the points in $\mathbb{R}^n \setminus \Omega$ (i.e., the points outside of Ω) that interact with the points inside of Ω . Below in (9), we define the specific interaction domain that is used in peridynamic modeling. From (7) or (8), we see that points in $\mathbf{y} \in \Omega_I$ affect the nonlocal operator

at points $\mathbf{x} \in \Omega$, thus the nomenclature “interaction domain.” Note that Ω_I is also the region on which the solution may be constrained.

For nonlocal models such as peridynamics, an important issue is the proper understanding of the nonlocal analog of boundary conditions specified for PDEs. In Du et al. (2012a), the notion of volume-constrained problems is discussed because in general, unlike the case of local PDEs, interactions occur at a distance. Of course, this does not exclude the case of classical local boundary conditions in special cases. For example, in recent studies, linear peridynamic models with a variable horizon $\varepsilon = \varepsilon(\mathbf{x})$ for $\mathbf{x} \in \Omega$ have also been studied Silling et al. (2014) and Tian and Du (2016). In such a case, a variable horizon $\varepsilon(\mathbf{x})$ is allowed to vanish as material points approach the boundary Γ of Ω , e.g., $\varepsilon(\mathbf{x}) \propto \text{dist}(\mathbf{x}, \Gamma)$ for points \mathbf{x} near the boundary Γ . In this case, displacement boundary conditions can be directly imposed on Γ based on the new trace theorem proved in Tian and Du (2016).

Concerning nonlocal problems defined on a bounded domain, we follow the discussion in Du et al. (2012a). First of all, one may define a notion of essential constraints and natural or variational conditions for nonlocal problems by drawing analogies to classical variational problems for PDEs that are defined with respect to suitable boundary conditions. Indeed, with a constant horizon, the most widely studied case is that of Dirichlet-type essential constraints, i.e., displacement constraints on the solution that are imposed on a constraint set Ω_I of nonzero measure. For example, we may impose the condition $\mathbf{u} = 0$ on Ω_I ; this is an essential condition that must be built explicitly into the variational principle. Typically, we have the values of \mathbf{u} specified on

$$\Omega_I = \{\mathbf{x} \in \mathbb{R}^n \setminus \Omega : \text{dist}(\mathbf{x}, \partial\Omega) < \varepsilon\} \quad (9)$$

which is a ε -collar surrounding Ω , where $\partial\Omega$ denotes the boundary of Ω and also the inner boundary of Ω_I .

The case of Neumann-type conditions that, for peridynamics, are the analog of the natural traction condition for the PDE case can take different forms. If Ω_I is an empty set, then \mathcal{L}_ε can be well-defined over all of $\Omega \cup \Omega_I = \Omega$ without the need to impose additional conditions on \mathbf{u} . In this case, the forcing function \mathbf{b} may represent either a soft or hard loading with the distinction only showing up through the dependence on the horizon parameter. Another interpretation that is more symbolically similar to the PDE case is to have a nonempty Ω_I in the definitions of the nonlocal operators as that given by (8) and imposing (7) on both Ω that represents the nonlocal equation and Ω_I that represents the domain of nonlocal constraints or nonlocal traction conditions. Again, viewed in terms of $\Omega \cup \Omega_I$, such a separation of the equation domain and the constraint domain becomes mathematically superfluous, although \mathbf{b} can be subject to different physical interpretations, i.e., in Ω it represents a body force, whereas in Ω_I it represents a nonlocal traction force.

Variational Formulation

Related to Eq. (7), we consider the nonlocal constrained value problem

$$\begin{cases} -\mathfrak{L}_\varepsilon \mathbf{u}(\mathbf{x}) = \mathbf{b}(\mathbf{x}) \quad \forall \mathbf{x} \in \Omega \\ \mathbf{u}(\mathbf{x}) = \mathbf{0} \quad \forall \mathbf{x} \in \Omega_I \end{cases} \tag{10}$$

which may be derived by minimizing the corresponding energy functional

$$E(\mathbf{u}) = \int_{\Omega \cup \Omega_I} \int_{\Omega \cup \Omega_I} \frac{1}{2} \omega_\varepsilon(|\mathbf{y} - \mathbf{x}|) (\mathfrak{D}^* \mathbf{u}(\mathbf{x}, \mathbf{y}))^2 d\mathbf{y} d\mathbf{x} - \int_{\Omega \cup \Omega_I} \mathbf{b}(\mathbf{x}) \mathbf{u}(\mathbf{x}) d\mathbf{x}. \tag{11}$$

It is helpful to state clearly what type of kernels are amenable to variational frameworks and treatment using the nonlocal calculus. For example, ω_ε is often assumed to satisfy

1. $\omega_\varepsilon(r) \geq 0 \quad \forall r \in (0, \varepsilon)$ (12)

2. $\text{supp}(\omega_\varepsilon) \subset B_\varepsilon(0), \quad \text{i.e., } \underline{\omega}_\varepsilon(r) = 0 \quad r \geq \varepsilon,$ (13)

3. $\int_{\mathbb{R}^n} \omega_\varepsilon(|\mathbf{x}|) d\mathbf{x} = m < \infty,$ (14)

where the normalization condition (14) on ω_ε implies that it has finite second-order moments, a condition that is necessary for well-defined elastic moduli (Silling 2000). It is also equivalent to the requirement that the energy is finite for a linear displacement field. A consequence is that for any square integrable vector-valued function that has square integrable first-order partial derivatives, the energy remains finite. We note that the nonnegativity assumption $\omega_\varepsilon(r) \geq 0$ can be relaxed (Mengesha and Du 2013).

The corresponding energy space V contains square integrable functions, i.e., functions in $L^2(\Omega \cup \Omega_I)$ having finite energy. This is an inner product space corresponding to the norm

$$\|\mathbf{u}\|_v^2 = |\mathbf{u}|_v^2 + \|\mathbf{u}\|_{L^2(\Omega \cup \Omega_I)}^2,$$

where the semi-norm $|\mathbf{u}|_v$ on V is defined as

$$|\mathbf{u}|_v^2 = \int_{\Omega \cup \Omega_I} \int_{\Omega \cup \Omega_I} \omega_\varepsilon(|\mathbf{x} - \mathbf{y}|) (\mathfrak{D}^* \mathbf{u})^2 d\mathbf{y} d\mathbf{x}.$$

The solution space $V_{c,\varepsilon}$ is the subspace of V consisting of functions satisfying the constraint $\mathbf{u}(\mathbf{x}) = \mathbf{0}$ on Ω_I . Various properties of the function spaces V and $V_{c,\varepsilon}$

can be studied in either L^2 spaces or fractional Sobolev spaces for specific kernels (Du et al. 2013a) and general abstract solution spaces (Mengesha and Du 2014a).

A variational formulation of (7) is given by the minimization problem

$$\min_{\mathbf{u} \in V_{c,\varepsilon}} E(\mathbf{u}).$$

The well-posedness of the variational problem is studied in Du et al. (2013b), Du and Zhou (2011), Mengesha and Du (2013, 2014a,b) in which additional details may be found about suitable interpretations of the operators of the nonlocal vector calculus, especially for cases involving non-integrable interaction kernels.

Note that if we have an integrable nonlocal interaction kernel, the space V and its dual V^* are both equivalent to the standard Hilbert space $L^2(\Omega \cup \Omega_I)$. In general, the nonlocal spaces may be larger than the $H^1(\Omega \cup \Omega_I)$ space so that they may be used to produce solutions having less regularity (Du et al. 2012a; Du and Zhou 2011) than that of solution of the analogous PDEs; indeed, this is a motivation for considering nonlocal mechanical models like peridynamics. As mentioned earlier, when $\varepsilon \rightarrow 0$, the nonlocal problems are closely related to classical local models. Techniques for investigating the local limit of nonlocal peridynamic models include Taylor expansions (Emmrich and Weckner 2007; Silling and Lehoucq 2008), Fourier transforms (Du et al. 2013a; Du and Zhou 2011), consistency of weak forms (Du et al. 2014), and the nonlocal calculus of variations (Mengesha and Du 2013, 2014a,b). For the linear bond-based peridynamic model, the local limit recovers the linear Navier elasticity model with a special Poisson ratio.

Discretization Schemes

There have been several approaches taken for the discretization of nonlocal models, including peridynamics. These include meshfree methods, finite difference methods, finite element methods, and numerical quadrature and collocation approaches. Mathematically rigorous studies of convergence properties of such discretizations include, for example, Tian and Du (2013), in which a number of finite difference approximations are given for one-dimensional scalar nonlocal peridynamic equations with their convergence proved; discrete maximum principles are also shown along with error estimates. For collocation schemes, rigorous studies include (Zhang et al. 2016a). For theoretical convergence results for finite element approximations, see Tian and Du (2013) and Zhou and Du (2010) for one-dimensional or special two-dimensional models; the same analysis is given in Tian and Du (2014) for general multidimensional linear peridynamic bond-based and state-based models using conforming finite element spaces defined on arbitrary grids. Nonconforming discontinuous Galerkin approximations have been discussed in Tian and Du (2015).

Quadrature rules for finite element approximations are considered in Zhang et al. (2016b).

Here, due to space limitations, we briefly review recent results about asymptotically compatible schemes.

Asymptotically Compatible Schemes

It is known in practice that to obtain consistency between nonlocal models and corresponding local PDE models, the mesh or quadrature point spacing may have to be reduced at a faster pace than the reduction of the horizon parameter (Bobaru et al. 2009; Bobaru and Duangpanya 2010). Otherwise, there could potentially be complications and inconsistent limiting solutions when the horizon parameter is coupled proportionally to the discretization parameter (Seleson et al. 2016; Tian and Du 2013, 2014). Asymptotically compatible (AC) schemes, formally introduced in Tian and Du (2014), are numerical discretization of nonlocal models that converge to nonlocal continuum models for fixed horizon parameter and to the local models as the horizon vanishes for both discrete schemes with a fixed numerical resolution and for continuum models with increasing numerical resolution.

Let $h > 0$ denote the meshing parameter (or particle spacing) and ε the horizon parameter or even a more generic model parameter. The implications of the AC property are illustrated in Fig. 1. There, u_ε denotes the solution of the continuous nonlocal problem with $\varepsilon > 0$, u_0 the solution of the corresponding continuous local problem, $u_{\varepsilon,h}$ the solution of the discretized nonlocal problem, and $u_{0,h}$ the solution of the discretized local problem.

As seen from detailed studies given in Tian and Du (2013), some popular discretization schemes of nonlocal peridynamics fail to be AC. In particular, if ε is taken to be proportional to h , then as $h \rightarrow 0$, piecewise constant conforming finite element solutions actually converge to the incorrect limit, similarly to those based on simple Riemann sum quadrature approximations to nonlocal operators. In Tian and

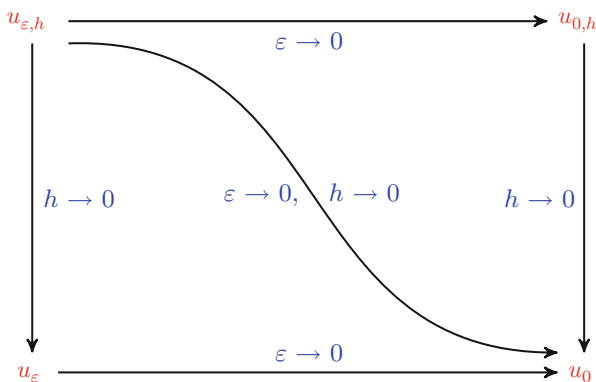


Fig. 1 A diagram for asymptotically compatible schemes Tian and Du (2014)

Du (2014) (see also Chen and Gunzburger 2011 for a computational illustration), it is shown that as long as the condition $h = o(\varepsilon)$ is met as $\varepsilon \rightarrow 0$, then we are able to obtain the correct local limit even for discontinuous piecewise constant finite element approximations when they are of the conforming type. Practically speaking, this implies that a mild growth of bandwidth is needed as the mesh is refined to recover the correct local limit for Riemann sum like quadratures or piecewise constant finite element schemes.

On the other hand, it is shown in Tian and Du (2014) (again, see also Chen and Gunzburger 2011 for computational illustrations) that all conforming Galerkin approximations of the nonlocal models containing continuous piecewise linear functions are automatically AC. This means that they can recover the correct local limit as long as both ε and h are diminishing, even if the nonlocality measure ε is reduced at a much faster pace than the mesh spacing h . Whereas the analysis is highly technical, an intuitive explanation is that even with h larger than ε , the nonlocal features are still encoded in the stiffness matrices to ensure the correct local limit due to the incorporation of higher-order (than constant) basis functions.

Naturally, schemes using higher-order basis functions tend to provide higher-order accuracy to the nonlocal problems as well. At the moment, the theory on AC schemes does not offer any estimate of the order of convergence with respect to the different coupling of h and ε . Preliminary numerical experiments in Tian and Du (2014) offer some insight about the balance of modeling and discretization errors, but additional theoretical analyses remain to be carried out.

Asymptotically compatible (AC) schemes, being either conforming Galerkin-type approximations of weak forms Chen and Gunzburger (2011), Tian and Du (2014), and Xu et al. (2016a,b), nonconforming DG approximations (Tian and Du 2015), or collocation- and quadrature-based approximations of strong forms (Du and Tian 2014; Seleson et al. 2016; Zhang et al. 2016a,b), offer the potential to solve the desirable models with different choices of parameters to gain efficiency and to avoid the pitfall of reaching inconsistent limits.

The framework of AC schemes is very general. For example, it has also been used to prove the convergence of numerical solutions of nonlocal diffusion equations to that of fractional diffusion equations Tian et al. (2016); see also D'Elia and Gunzburger (2013). The theory also guided the development of nonconforming DG approximations (Tian and Du 2015) for peridynamic models with non-integrable interaction kernels. The framework can also be applicable to numerical studies of nonlocal models of convection and diffusion (Tian et al. 2015, 2017) that were motivated by the work on nonlocal convection (Du et al. 2014) and nonlinear nonlocal hyperbolic conservation laws (Du and Huang 2017; Du et al. 2017a) that improved the model studied in Du et al. (2012b). Furthermore, such a concept can also be used in the study of nonlocal gradient recovery for the numerical solution of nonlocal models (Du et al. 2016).

Conditioning

A good theoretical understanding of the conditioning of stiffness matrices for nonlocal problems is alluded to in the Fourier analysis of the point spectrum for

the nonlocal operators as given in Du and Zhou (2016) and Zhou and Du (2010); see Aksoylu and Mengesha (2010), Aksoylu and Parks (2011), Aksoylu and Unlu (2013), Du et al. (2012a), and Seleson et al. (2009) for additional discussions. In comparison with a typical local diffusion model that yields a condition number of $O(h^{-2})$ for a discretization with meshing parameter h , the corresponding nonlocal models have their condition numbers dependent on both ε and h in general. For example, Du and Zhou (2016) provided sharp lower and upper bounds for finite element discretization of a nonlocal diffusion operator based on a quasi-uniform regular triangulation. For effective algebraic solvers of the resulting linear system, we refer to studies on the use of Toeplitz (Wang and Tian 2012) and multigrid (Du and Zhou 2016) solvers.

Adaptivity

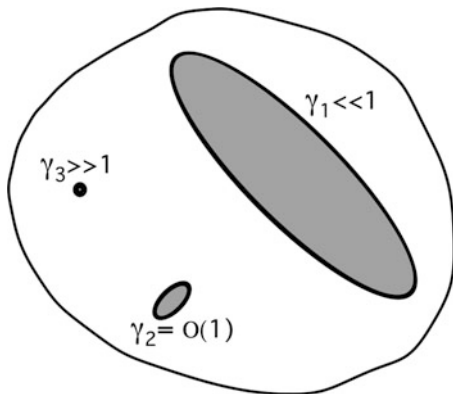
Due to, e.g., reduced sparsity, nonlocal models such as peridynamics generally incur greater computational costs than do their local PDE-based counterparts. Thus, designing effective adaptive methodologies is important. The paper Du et al. (2013c) contains an a posteriori error analysis of conforming finite element methods for solving linear nonlocal diffusion and peridynamic mechanic models. A general abstract framework is developed for a posteriori error analysis of nonlocal volume-constrained problems for scalar equations. The reliability and efficiency of the estimators are proved, and relations between nonlocal and classical local a posteriori error estimation are also studied. In Du et al. (2013d), a convergent adaptive finite element algorithm for the numerical solution of scalar nonlocal models is developed. For problems involving certain non-integrable kernel functions, the convergence of the adaptive algorithm is rigorously derived with the help of several basic ingredients, such as an upper bound on the estimator, the estimator reduction, and the orthogonality property. How these estimators and methods work in the local limit and for general time-dependent and nonlinear peridynamic models is under current investigation.

Peridynamics as a Multiscale Mono-model

Classical PDE models for mechanics such as the Navier equations for elasticity do not feature a length scale other than those related to material properties. On the other hand, peridynamics features the horizon ε that determines the extent of nonlocal interactions, even in the case of homogeneous, isotropic materials. The appearance of a length scale implies that the nature of the peridynamic model changes as the “viewing window” through which one looks at the model changes in size relative to ε . Here, we give a brief discussion of this observation.

Let $\widehat{\Omega} \subset \Omega$. Because nonlocal interactions only occur between points within a distance ε from each other, a point $\mathbf{x} \in \widehat{\Omega}$ only interacts with points $\mathbf{y} \in \widehat{\Omega} \cup \widehat{\Omega}_\varepsilon$, where $\widehat{\Omega}_\varepsilon = \{\mathbf{y} \in \Omega \cup \Omega_I \setminus \widehat{\Omega} \mid |\mathbf{y} - \mathbf{x}| \leq \varepsilon\}$. Thus, $\widehat{\Omega}_\varepsilon$ is the interaction domain

Fig. 2 The gray areas illustrate three regions Ω_i having different length scales. The black regions surrounding the gray regions are the interaction regions $\Omega_{i,\varepsilon}$, all of which have the same thickness ε



corresponding to $\widehat{\Omega}$. Note that $\widehat{\Omega}_\varepsilon$ is a collar surrounding $\widehat{\Omega}$. We denote by $|\widehat{\Omega}|$ and $|\widehat{\Omega}_\varepsilon|$ the volumes of Ω and Ω_ε , respectively.

Let $d = \text{diam}(\widehat{\Omega})$ and assume that $\widehat{\Omega}$ is “nicely shaped” so that, in \mathbb{R}^3 , $|\widehat{\Omega}| = O(d^3)$. Then, $|\widehat{\Omega}_\varepsilon| = O(d^2\varepsilon + d\varepsilon^2 + \varepsilon^3)$. We have that the *volume ratio* $\gamma = |\widehat{\Omega}_\varepsilon|/|\widehat{\Omega}| = O(\varepsilon/d + \varepsilon^2/d^2 + \varepsilon^3/d^3)$ is an indicator of the relative nonlocality for the subdomain $\widehat{\Omega}$. If ε is a fixed material parameter, γ decreases (resp. increases) as $|\widehat{\Omega}|$ increases (resp. decreases). Because there are no self-interactions (a consequence of Newton’s third law), the net internal force on $\widehat{\Omega}$ due to the action of points in $\widehat{\Omega}$ is zero so that the internal force on $\widehat{\Omega}$ is solely due to interactions with points in $\widehat{\Omega}_\varepsilon$.

We introduce three subdomains $\widehat{\Omega}_i$ with diameters d_i , interaction domains $\widehat{\Omega}_{i,\varepsilon}$, and volume ratios γ_i , $i = 1, 2, 3$; see Fig. 2. We consider three cases:

- $d_1 \gg \varepsilon$ so that $\gamma_1 = O(\varepsilon/d_1) \ll 1$ so that the internal force acting on $\widehat{\Omega}_1$ is *local*, emanating from the very thin (relative to the size of $\widehat{\Omega}_1$) layer $\widehat{\Omega}_{1,\varepsilon}$ of material points surrounding $\widehat{\Omega}$;
- $d_2 = O(\varepsilon)$ so that $\gamma_2 = O(1)$ and the internal force acting on $\widehat{\Omega}_2$ is *nonlocal*, emanating from the layer $\widehat{\Omega}_{2,\varepsilon}$ of material points having roughly the same size as $\widehat{\Omega}_2$;
- $d_3 \ll \varepsilon$ so that $\gamma_3 = O(\varepsilon^3/d_3^3) \gg 1$ so that the internal force acting on $\widehat{\Omega}_3$ is *nonlocal*, emanating from the very thick (relative to the size of $\widehat{\Omega}_3$) layer $\widehat{\Omega}_{3,\varepsilon}$ of material points surrounding $\widehat{\Omega}_3$.

Thus, depending on the scale one is operating in, peridynamics can behave as either a local or nonlocal material model so that, clearly, *peridynamics is by itself a multiscale model for material mechanics*, hence our characterization of peridynamics as a *multiscale mono-model* for mechanics. This observation has important implications for computations which we next explore.

A Multiscale Finite Element Implementation of Peridynamics

The multiscale nature of peridynamic models resulting from the horizon ε motivates a multiscale finite element implementation of peridynamics. Suppose h is a typical grid size in a discretization of a PDE model for solid mechanics in case the solution, i.e., the displacement, is known to be smooth; for example, if one is interested in achieving an L^2 error of $O(\delta)$ and one uses continuous piecewise linear finite element approximations, then h would be chosen to be of $O(\delta^{1/2})$. Even in cases where the solution has discontinuities, one can safely use the PDE model in regions away from the discontinuity, i.e., in regions where the displacement is smooth. However, near the location of discontinuities in the displacement, the PDE model breaks down, and one instead uses a peridynamic model. However, because the locations of those discontinuities are, in general, not known beforehand, one has to be able to detect where they occur. One also has to devise a strategy for coupling the peridynamic and PDE models. The overall goal of the implementation is to achieve, e.g., for piecewise linear approximations $O(h^2)$ accuracy even in the present of discontinuities.

Based on the above discussions, a multiscale implementation of the peridynamic model in one dimension is presented in Chen and Gunzburger (2011) and Xu et al. (2016a) and in two dimensions in Xu et al. (2016b). One starts with a choice for a horizon parameter ε and for a bulk grid size h , that is, what one would use should one be solving a PDE model for problems with smooth solutions. Then, the multiscale implementation includes the following components:

1. Detection of elements that contain a discontinuity in the displacement. A simple means for doing this is to examine the size of the difference in the displacements at the nodes of a triangle. More sophisticated means for this step are also considered in Xu et al. (2016b).
2. Refinement of the grid as necessary near the discontinuities. Meeting the goal of refinement requires the use of very small elements so that refinement should not result in isotropic elements containing the discontinuity because this would result in a huge number of such elements. Instead, refinement should result in highly anisotropic elements with long sides aligned with a discontinuity and very short sides crossing a discontinuity.
3. Use of discontinuous Galerkin methods (DGMs) for peridynamics in regions containing the discontinuity. The kernel functions used in peridynamics are such that such methods are conforming. In particular, unlike the PDE case, one need not worry about enforcing continuity of any kind across element edges; nonlocality takes care of a weak enforcement of such properties.
4. Use of continuous Galerkin methods (CGMs) for peridynamics in regions neighboring but not containing the discontinuity. The use of CGMs wherever possible results in a reduction in the number of degrees of freedom compared to using DGMs.

5. Use of CGMs for PDEs if sufficiently far away from the discontinuity. In regions where the displacement is smooth, PDEs provide adequate modeling and CGMs provide conforming discretizations.
6. Use of quadrature rules that can be applied for any combination of h and ε . Such quadrature rules on triangles and regions defined as the intersection of triangles with the interaction disc of radius ε are developed in Xu et al. (2016b) that are applicable to the integrable kernels appearing in two-dimensional peridynamic models.

Generically, one would choose $h \gg \varepsilon$ because then the discretized peridynamic model essentially reduces to a local model whenever the local grid size is of $O(h)$. Specifically, if one uses the same asymptotically compatible continuous finite element space and the same basis to discretize both the nonlocal peridynamic model and the corresponding local PDE model, it can be shown that the difference between the entries of the peridynamic stiffness and those of PDE stiffness matrix is of $O(\varepsilon)$. This effect greatly facilitates the coupling of a peridynamic model to a PDE model; in fact, the coupling is seamless. On the other hand, we would also choose $\varepsilon > \underline{h}$, where \underline{h} denotes the grid size one would use for elements in which solution discontinuities occur. With $\varepsilon > \underline{h}$, the peridynamic model is nonlocal. To meet the goal of achieving $O(h^2)$ accuracy, one should choose $\underline{h} = O(h^4)$ because in any element containing a solution discontinuity whose position is not known exactly, the best accuracy one can achieve is $O(\underline{h}^{1/2})$ regardless of what degree polynomial one uses. It is shown in Chen and Gunzburger (2011) and Xu et al. (2016a,b) that the fine-to-coarse grid transition can, at least for kernels that allow for discontinuous displacements and for DGMs, be abrupt, i.e., there is no need for transition layer within which the element size gradually grows from small to fine. DGMs need only be used in elements that contain the discontinuity.

We illustrate the multiscale implementation strategy for an exact solution having a jump discontinuity across a circle of radius 0.5 centered within the square $[0, 1]^2$; the fixed horizon $\varepsilon = 0.01$. Starting with a uniform mesh of size h , the target grid size for regions where the solution is smooth, an adaptive grid refinement strategy that first detects elements containing the discontinuity and which ultimately constructs a grid consisting of $O(h)$ size elements except for highly anisotropic elements (the elements connecting the red nodes in Fig. 3 which is for $h = 1/16$) along the discontinuity that are of size $O(h)$ parallel to the discontinuity and of size $O(h^4)$ across the discontinuity, the actual thickness of those elements is of $O(10^{-5})$ so that they are not visible. In Fig. 3, DGM and peridynamics are used in elements with a red node, CGM and peridynamics are used in elements having both red and blue nodes, and CGM with PDE are used in the remaining elements. Note that the highly anisotropic straight-sided elements cannot, in general, contain the discontinuity throughout the element. The construction process merely insures that the discontinuity passes through the shortest side of those elements. We have empirically found that doing this is enough to preserve optimal accuracy.

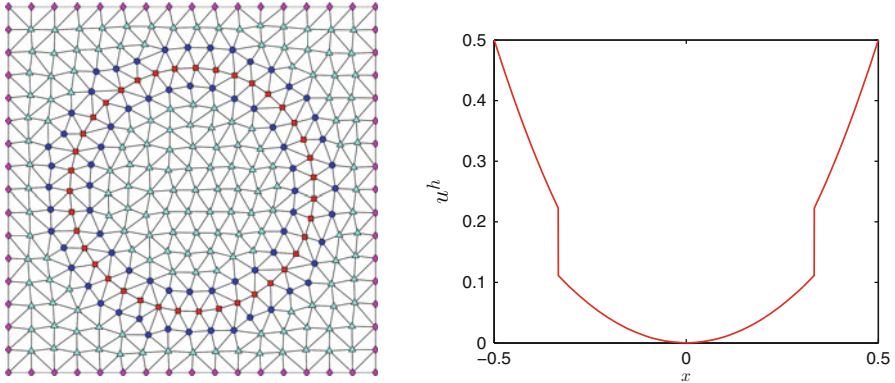


Fig. 3 For a problem with a discontinuous solution across a circle, the resulting adaptively refined mesh (left) and a cross-section of approximate solution (right). Note that all elements are of size h excepting the thin elements connecting the red nodes which are not visible because they have one side of length $O(h^4)$

As a result of these implementation choices, the number of degrees of freedom needed is only slightly greater than that for a finite element method on an $O(h)$ size grid for the PDE. The overhead incurred through the use of a nonlocal model to treat discontinuous behaviors is due to the use of such models in a layer of highly anisotropic elements surrounding the discontinuity and the use of DGMs in those elements. Because, in two dimensions, the number of such elements grows linearly, whereas the total number of elements grows quadratically, the percentage overhead actually reduces as h decreases.

The implementation of Chen and Gunzburger (2011) and Xu et al. (2016a,b) clearly illustrates the multiscale mono-model nature of peridynamics but is still a far cry from being a useful tool; much work needs to be done to achieve the latter, with the most daunting tasks remaining being extensions to three dimensions and the treatment of problems with propagating, i.e., moving, discontinuities.

Control, Identification, and Obstacle Problems

In this section, we very briefly touch upon some other recent developments for nonlocal models that go beyond standard simulation settings.

Nonlocal Optimal Control Problems

The ingredients of an optimal control problem are state and control variables, a state equation that relates the state and control variables, and a cost or objective functional which depends on the state and control. Then, the goal of the optimal

control problem is to determine states and controls that minimize the cost functional, subject to the state equation being satisfied.

We consider state equations of the form

$$\begin{cases} -\mathcal{L}u(\mathbf{x}) := 2 \int_{B_\varepsilon(\mathbf{x})} (u(\mathbf{y}) - u(\mathbf{x})) \beta(\mathbf{x}, \mathbf{y}) d\mathbf{y} = f & \text{for } \mathbf{x} \in \Omega \\ u = g & \text{for } \mathbf{x} \in \Omega_I, \end{cases} \quad (15)$$

where $u(\mathbf{x})$ denotes the state, $f(\mathbf{x})$ the control, $g(\mathbf{x})$ a given function, and $\beta(\mathbf{x}, \mathbf{y})$ a symmetric kernel, i.e., $\beta(\mathbf{x}, \mathbf{y}) = \beta(\mathbf{y}, \mathbf{x})$. We consider the matching functional

$$J(u, f) := \frac{1}{2} \int_{\Omega} (u(\mathbf{x}) - \widehat{u}(\mathbf{x}))^2 d\mathbf{x} + \frac{\sigma}{2} \int_{\Omega} (f(\mathbf{x}))^2 d\mathbf{x}, \quad (16)$$

where $\widehat{u}(\mathbf{x})$ is a given target function and $\sigma > 0$ a given parameter. The first term in (16) is the goal of control, i.e., to match (in an L^2 sense for the functional $J(u, f)$) the solution $u(\mathbf{x})$ of (16) as well as possible to the given target function $\widehat{u}(\mathbf{x})$. Then, the optimal control problem is to

seek an optimal state $u^*(\mathbf{x})$ and an optimal control $f^*(\mathbf{x})$ such that the functional (16) is minimized subject to (15) being satisfied.

Of course, a precise statement of this problem requires the choice of appropriate function spaces for u and f ; these are defined in D’Elia and Gunzburger (2014). There, the existence and uniqueness of the optimal control and state in appropriate function spaces are also proved for several kernels relevant to peridynamics or anomalous diffusion.

The optimal state $u^*(\mathbf{x})$ is determined from the optimality system

$$\begin{cases} -\mathcal{L}u^* = -\frac{1}{\sigma} w^* & \text{for } \mathbf{x} \in \Omega \\ u^* = g & \text{for } \mathbf{x} \in \Omega_I \end{cases}$$

and

$$\begin{cases} -\mathcal{L}w^* = (u^* - \widehat{u}) & \text{for } \mathbf{x} \in \Omega \\ w^* = 0 & \text{for } \mathbf{x} \in \Omega_I. \end{cases} \quad (17)$$

The adjoint function $w^*(\mathbf{x})$ is a Lagrange multiplier function that is introduced to enforce the constraint (15) imposed when minimizing the functional (16).

After the optimal state $u^*(\mathbf{x})$ and optimal adjoint state $w^*(\mathbf{x})$ are obtained, the optimal control $f^*(\mathbf{x})$ is given by

$$f^* = -\frac{1}{\sigma} w^* \quad \text{for } \mathbf{x} \in \Omega.$$

Also proved in D’Elia and Gunzburger (2014) is the convergence, as $\varepsilon \rightarrow 0$, of solutions of the nonlocal control problem to solutions of the local control problem for which the state equation (15) is replaced by the PDE problem

$$\begin{cases} -\Delta u_\ell = f_\ell & \text{for } \mathbf{x} \in \Omega \\ u_\ell = g_\ell & \text{for } \mathbf{x} \in \partial\Omega, \end{cases} \tag{18}$$

where $u_\ell(\mathbf{x})$ and $f_\ell(\mathbf{x})$ denote the local PDE solution and control, respectively, and $\partial\Omega$ denotes the boundary of Ω . The given Dirichlet data $g_\ell(\mathbf{x})$ is determined by evaluating the nonlocal data $g(\mathbf{x})$ on $\partial\Omega$; this, of course, requires $g(\mathbf{x})$ for $\mathbf{x} \in \Omega_I$ to be more regular than that needed for the nonlocal problem to be well posed. In addition, in D’Elia and Gunzburger (2014), the convergence of solutions of finite-dimensional discretizations of the nonlocal optimality system, including finite element discretizations, is proved, and optimal error estimates are derived, again for several kernels relevant to peridynamics and anomalous diffusion.

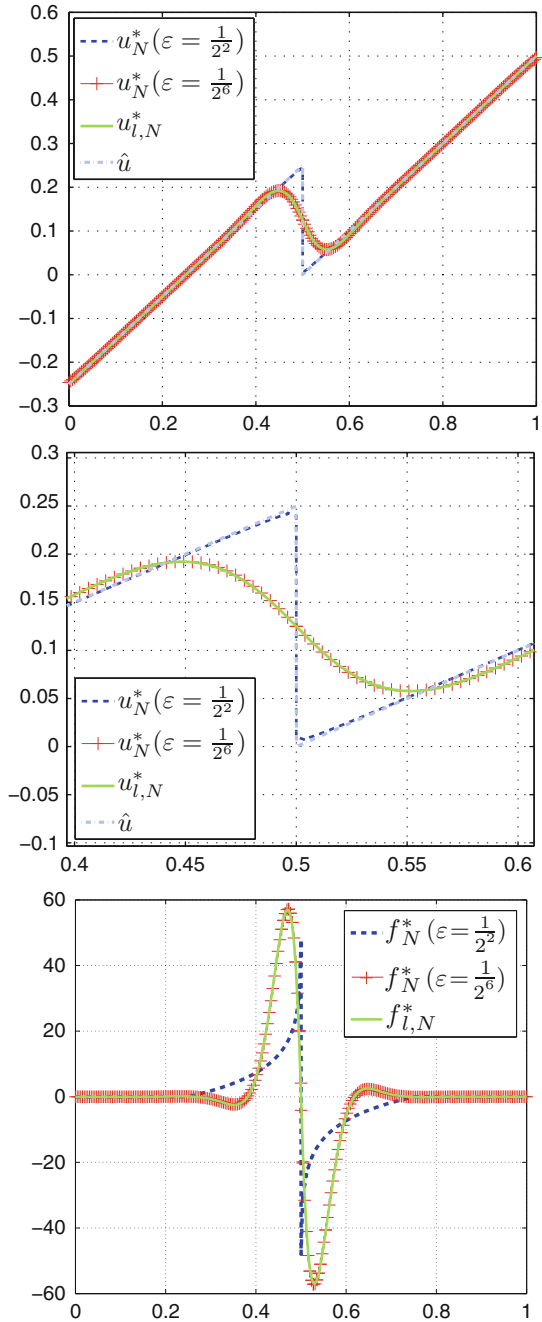
As an example, we consider a piecewise linear target function having a jump discontinuity; see the light blue curve in Fig. 4. In that figure, also plotted are the target functional \widehat{u} and, for two values of ε , the finite element approximations u_N^* of the optimal state and f_N^* of the optimal control. We observe that, away from the point $x = 0.5$ at which the target function is discontinuous, the local and nonlocal approximations show a very good match to the target function. However, we observe that because the local optimal state is differentiable, its approximation cannot match the target function near $x = 0.5$. Also, because the nonlocal optimal state converges to the local optimal state as $\varepsilon \rightarrow 0$, we see that for the smaller value of ε , the nonlocal and local optimal states are in close agreement so that, in this case, the nonlocal optimal state also fails to match the target function near the discontinuity. This is to be expected because we have already stated that the solution of the nonlocal problem converges to that of the local problem as $\varepsilon \rightarrow 0$. However, for the larger value of ε for which nonlocal effects are more pronounced, we observe a difference between the approximation of the local optimal state (which does not match the target well) and the approximation of the nonlocal optimal state (which do match well), being more flexible with respect to the discontinuous behavior of the target function.

Nonlocal Parameter Identification Problems

We next consider the problem of identifying the diffusivity coefficient function $\vartheta(\mathbf{x}, \mathbf{y})$ in the nonlocal equation

$$\begin{cases} 2 \int_{B_\varepsilon(\mathbf{x})} (u(\mathbf{y}) - u(\mathbf{x})) \vartheta(\mathbf{x}, \mathbf{y}) \beta(\mathbf{x}, \mathbf{y}) d\mathbf{y} = f(\mathbf{x}) & \text{for } \mathbf{x} \in \Omega \\ u(\mathbf{x}) = g(\mathbf{x}) & \text{for } \mathbf{x} \in \Omega_I, \end{cases} \tag{19}$$

Fig. 4 The target function and finite element approximations of the optimal state (left) and control (right) for the nonlocal and local control problems. The middle plot is a zoom-in of the left plot in the vicinity of the discontinuity in the target function



where the kernel $\beta(\mathbf{x}, \mathbf{y}): \Omega \times \Omega \rightarrow \mathbb{R}$ is a nonnegative symmetric mapping and $f(\mathbf{x})$ and $g(\mathbf{x})$ are given functions. We assume that the diffusivity function satisfies

$$0 < \vartheta_0 \leq \vartheta(\mathbf{x}, \mathbf{y}) \leq \vartheta_1 < \infty \quad (20)$$

for some constants ϑ_0 and ϑ_1 . Additional technical assumptions on $\vartheta(\mathbf{x}, \mathbf{y})$ are needed for the analysis; see D'Elia and Gunzburger (2016).

Our approach is to

seek an optimal state $u^*(\mathbf{x})$ and an optimal diffusivity $\vartheta^*(\mathbf{x}, \mathbf{y})$ such that the functional $K(u, \vartheta)$ is minimized subject to (19) and (20) being satisfied,

where, for a given target function $\widehat{u}(\mathbf{x})$, we have the functional

$$K(u, \vartheta) := \frac{1}{2} \int_{\Omega} (u(\mathbf{x}) - \widehat{u}(\mathbf{x}))^2 d\mathbf{x}. \quad (21)$$

Note the absence of a regularization term in (21) compared to the functional (16); that term is not needed in the analysis because of (20) and the other technical assumptions explicitly enforced on candidate solutions.

In D'Elia and Gunzburger (2016), the existence of a solution of this parameter identification problem is proved. In general, the solution is not unique. This is not due to the nonlocality; lack of uniqueness also holds for the corresponding local PDE parameter identification problem. In that paper, an optimality system whose solution provides the optimal state and optimal diffusivity coefficient is also derived. For the sake of brevity, we do not discuss this here other than to say that in addition to the state equation (19) and an adjoint equation similar to (17), complementary conditions are present that enforce the constraints in (20). Also in D'Elia and Gunzburger (2016), finite element discretizations of the optimality system are defined, and convergence proofs for the discrete solution are provided, including a priori error estimates for the approximate states and coefficients.

As an illustration, we assume the form $\vartheta(x, y) = \vartheta(\frac{x+y}{2})$ for the coefficient and then choose the particular piecewise constant coefficient

$$\vartheta_C(z) = \begin{cases} 1 & z \in (0, 0.2) \\ 0.1 & z \in (0.2, 0.6) \\ 1 & z \in (0.6, 1). \end{cases} \quad (22)$$

Piecewise constant coefficients often arise in practice; for example, in subsurface flows, disjoint portions of the subsurface may consist of rock, clay, or sand. Using that coefficient, we solve the state equation (19) with $f(x) = 5$ on $\Omega = (0, 1)$, $u(x) = 0$ on $\Omega_I = (-\varepsilon, 0) \cup (1, 1 + \varepsilon)$, and a peridynamic-like kernel to produce the manufactured target function $\widehat{u}(x)$ for the functional (21); in actuality, we solve for a very fine grid finite element approximation of $\widehat{u}(x)$. Having done this, we pretend we do not know either the state or coefficient but try to identify them by solving the optimality system. Note that the diffusivity coefficient we try to identify

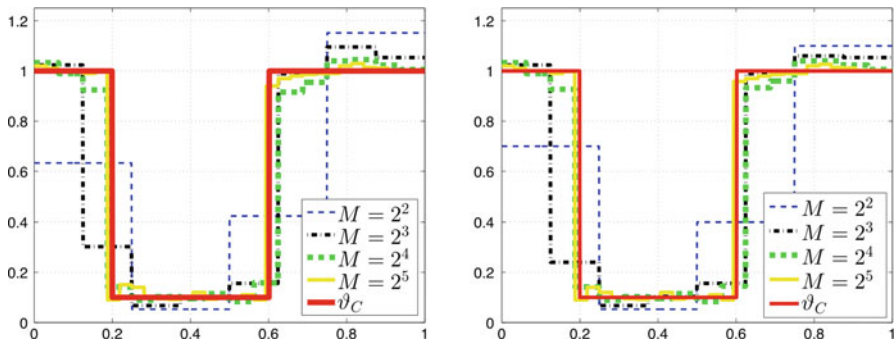


Fig. 5 For different values of M and for $N = 2^{12}$, the approximation ϑ_M^* of the optimal coefficient for $\varepsilon = 2^{-9}$ (left) and 2^{-4} (right). Also plotted is the target coefficient ϑ_C

is discontinuous so that a local PDE-based parameter identification problem would do a poor job of matching that coefficient.

Computational results are shown in Fig. 5. In that figure, M denotes the number of degrees of freedom used in a piecewise constant approximation of ϑ . Note that M is different from the number of degrees N used for a continuous piecewise linear approximation of u . In Fig. 5, we use a large value for N so as to minimize the effects of inaccuracies present in the approximate solution of the constraint equation (19). From Fig. 5, one observes that as M increases, the approximation ϑ^* resulting from the optimization process converges to the presumably unknown exact coefficient ϑ_C . We again note that nonlocality allows for much better match than is possible with PDE-based models.

Nonlocal Obstacle Problems

We consider the nonlocal obstacle problem given by the nonlocal equation (15) with $g = 0$ for which the solution $u(\mathbf{x})$ is subject to the constraint

$$u(\mathbf{x}) \geq \psi(\mathbf{x}) \quad \text{for } \mathbf{x} \in \Omega, \tag{23}$$

where $\psi(\mathbf{x})$ is a given function. This problem is equivalent to the variational problem

$$\min_u I[u] \quad \text{subject to } u \geq \psi \text{ in } \Omega \quad \text{and } u = 0 \text{ on } \Omega_I \tag{24}$$

for the functional

$$I[u] := \frac{1}{2} \int_{\Omega \cup \Omega_I} \int_{\Omega \cup \Omega_I} \beta(\mathbf{x}, \mathbf{y})(u(\mathbf{x}) - u(\mathbf{y}))^2 d\mathbf{y} d\mathbf{x} - \int_{\Omega} u(\mathbf{x}) f(\mathbf{x}) d\mathbf{x}.$$

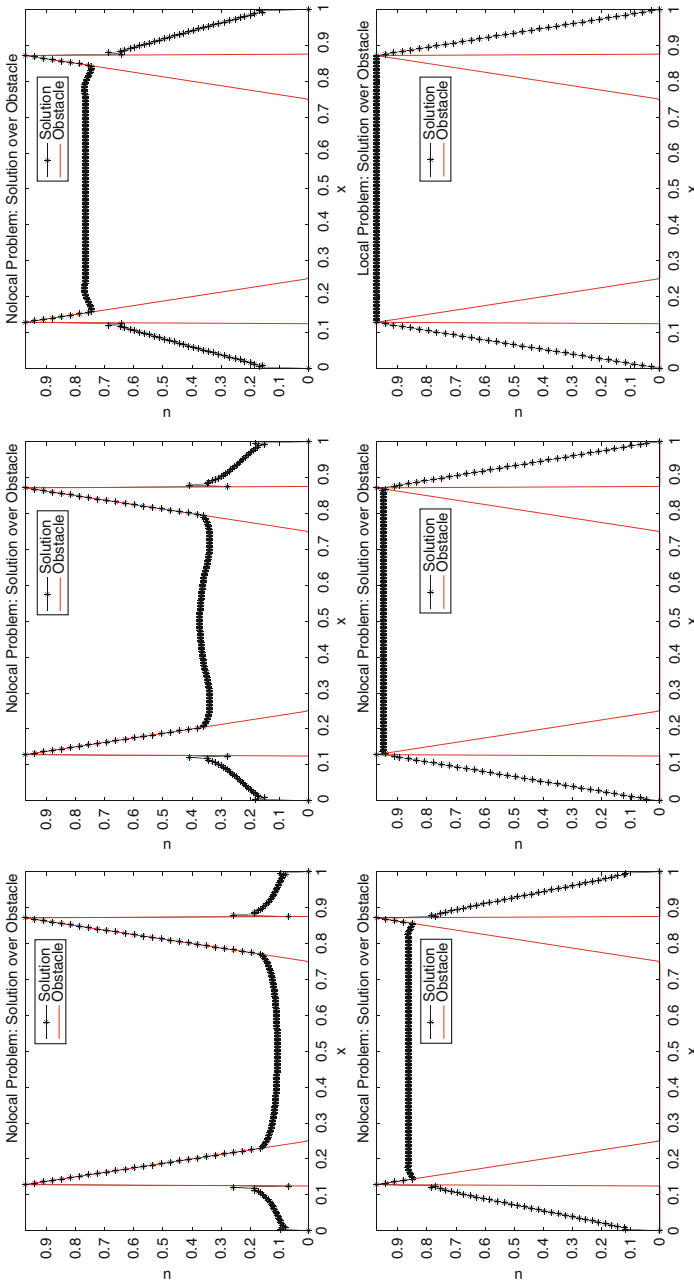


Fig. 6 Approximate solutions of the local obstacle problem (bottom-right) and of the nonlocal obstacle problem for $\epsilon = 2, 0.5, 0.1, 0.05, \text{ and } 0.01$ (left to right and top to bottom)

In Guan and Gunzburger (2017), the existence and uniqueness of solutions of the nonlocal obstacle problem (24) is proved for integrable and certain classes of non-integrable kernels. There, it is also proved that finite element discretizations are well posed as is its convergence to the solution of the continuous problem. In addition, the convergence as $\varepsilon \rightarrow 0$ of solutions of the nonlocal obstacle problem to that of the local obstacle problem is also proved under certain assumptions.

The corresponding local obstacle problem is given by Eq. (18) subject to the constraint (23). Well-posedness proofs for this problem require more smoothness for the obstacle than is required for the nonlocal obstacle problem. In particular, for the local case, obstacles with jump discontinuities are not allowable, whereas for the nonlocal case, for integrable and certain classes of non-integrable kernels with relatively weak singularities, such obstacles fall within the purview of the well-posedness and finite element analyses of Guan and Gunzburger (2017).

There is also a very significant difference in the solutions of the local and nonlocal obstacle problems, especially for non-smooth obstacles. An illustration is provided by Fig. 6 which features a piecewise linear obstacle having two jump discontinuities. The solution of the obstacle problem is given for five values of ε as is the solution of the local obstacle problem. We see that, as expected, for small values of ε , the nonlocal solution is close to the local ones, but for larger values, there is a significant difference, with the nonlocal solution being able to conform much better to the shape of the obstacle, even on its vertical edges. The local solution can only stretch over the peaks of the obstacle and completely ignores the shape of the obstacle below those peaks.

Acknowledgements MD: supported by the Collaboratory on Mathematics for Mesoscopic Modeling of Materials (CM4). Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the US Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525.

QD: supported in part by the US NSF grant DMS-1558744, the AFOSR MURI Center for Material Failure Prediction Through Peridynamics, and the OSD/ARO/MURI W911NF-15-1-0562 on Fractional PDEs for Conservation Laws and Beyond: Theory, Numerics and Applications.

MG: supported by the US NSF grant DMS-1315259, US Department of Energy Office of Science grant DE-SC0009324, US Air Force Office of Scientific Research grant FA9550-15-1-0001, and DARPA Equips program through the Oak Ridge National Laboratory.

References

- B. Aksoylyu, T. Mengesha, Results on nonlocal boundary value problems. *Numer. Func. Anal. Optim.* **31**, 1301–1317 (2010)
- B. Aksoylyu, M. Parks, Variational theory and domain decomposition for nonlocal problems. *Appl. Math. Comp.* **217**, 6498–6515 (2011)
- B. Aksoylyu, Z. Unlu, Conditioning analysis of nonlocal integral operators in fractional Sobolev spaces. *SIAM J. Numer. Anal.* **52**(2), 653–677 (2014). <https://doi.org/10.1137/13092407X>
- E. Askari, F. Bobaru, R. Lehoucq, M. Parks, S. Silling, O. Weckner, Peridynamics for multiscale materials modeling. *J. Phys. Conf. Ser.* **125**, 012078 (2008)

- F. Bobaru, M. Duangpanya, The peridynamic formulation for transient heat conduction. *Int. J. Heat Mass Transf.* **53**, 4047–4059 (2010)
- F. Bobaru, S. Silling, Peridynamic 3D problems of nanofiber networks and carbon nanotube-reinforced composites, in *Materials and Design: Proceeding of International Conference on Numerical Methods in Industrial Forming Processes* (American Institute of Physics, 2004), pp. 1565–1570
- F. Bobaru, S. Silling, H. Jiang, Peridynamic fracture and damage modeling of membranes and nanofiber networks, in *Proceeding of XI International Conference on Fracture*, Turin, 2005, pp. 1–6
- F. Bobaru, M. Yang, L. Alves, S. Silling, E. Askari, J. Xu, Convergence, adaptive refinement, and scaling in 1d peridynamics. *Inter. J. Numer. Meth. Engrg.* **77**, 852–877 (2009)
- X. Chen, M. Gunzburger, Continuous and discontinuous finite element methods for a peridynamics model of mechanics. *Comput. Meth. Appl. Mech. Engrg.* **200**, 1237–1250 (2011) with X. Chen.
- M. D'Elia, M. Gunzburger, The fractional Laplacian operator on bounded domains as a special case of the nonlocal diffusion operator. *Comput. Math. Appl.* **66**, 1245–1260 (2013)
- M. D'Elia, M. Gunzburger, Optimal distributed control of nonlocal steady diffusion problems. *SIAM J. Cont. Optim.* **52**, 243–273 (2014)
- M. D'Elia, M. Gunzburger, Identification of the diffusion parameter in nonlocal steady diffusion problems. *Appl. Math. Optim.* **73**, 227–249 (2016)
- Q. Du, Nonlocal calculus of variations and well-posedness of peridynamics, in *Handbook of Peridynamic Modeling*, chapter 3 (CRC Press, Boca Raton, 2016a), pp. 61–86
- Q. Du, Local limits and asymptotically compatible discretizations, in *Handbook of Peridynamic Modeling*, chapter 4 (CRC Press, Boca Raton, 2016b), pp. 87–108
- Q. Du, Z. Huang, Numerical solution of a scalar one-dimensional monotonicity-preserving nonlocal nonlinear conservation law. *J. Math. Res. Appl.* **37**, 1–18 (2017)
- Q. Du, X. Tian, Asymptotically compatible schemes for peridynamics based on numerical quadratures, in *Proceedings of the ASME 2014 International Mechanical Engineering Congress and Exposition IMECE 2014–39620*, 2014
- Q. Du, K. Zhou, Mathematical analysis for the peridynamic nonlocal continuum theory. *Math. Model. Numer. Anal.* **45**, 217–234 (2011)
- Q. Du, Z. Zhou, Multigrid finite element method for nonlocal diffusion equations with a fractional kernel (2016, preprint)
- Q. Du, M. Gunzburger, R. Lehoucq, K. Zhou, Analysis and approximation of nonlocal diffusion problems with volume constraints. *SIAM Rev.* **56**, 676–696 (2012a)
- Q. Du, J. Kamm, R. Lehoucq, M. Parks, A new approach to nonlocal nonlinear conservation laws. *SIAM J. Appl. Math.* **72**, 464–487 (2012b)
- Q. Du, M. Gunzburger, R. Lehoucq, K. Zhou, A nonlocal vector calculus, nonlocal volume-constrained problems, and nonlocal balance laws. *Math. Mod. Meth. Appl. Sci.* **23**, 493–540 (2013a)
- Q. Du, M. Gunzburger, R. Lehoucq, K. Zhou, Analysis of the volume-constrained peridynamic Navier equation of linear elasticity. *J. Elast.* **113**, 193–217, (2013b)
- Q. Du, L. Ju, L. Tian, K. Zhou, A posteriori error analysis of finite element method for linear nonlocal diffusion and peridynamic models. *Math. Comput.* **82**, 1889–1922 (2013c)
- Q. Du, L. Tian, X. Zhao, A convergent adaptive finite element algorithm for nonlocal diffusion and peridynamic models. *SIAM J. Numer. Anal.* **51**, 1211–1234 (2013d)
- Q. Du, Z. Huang, R. Lehoucq, Nonlocal convection-diffusion volume-constrained problems and jump processes. *Disc. Cont. Dyn. Sys. B* **19**, 373–389 (2014)
- Q. Du, Y. Tao, X. Tian, J. Yang, Robust a posteriori stress analysis for approximations of nonlocal models via nonlocal gradient. *Comput. Meth. Appl. Mech. Eng.* **310**, 605–627 (2016)
- Q. Du, Z. Huang, P. Lefloch, Nonlocal conservation laws. I. A new class of monotonicity-preserving models. *SIAM J. Numer. Anal.* **55**(5), 2465–2489 (2017)
- E. Emmrich, O. Weckner, On the well-posedness of the linear peridynamic model and its convergence towards the Navier equation of linear elasticity. *Commun. Math. Sci.* **5**, 851–864 (2007)

- W. Gerstle, N. Sau, Peridynamic modeling of concrete structures, in *Proceeding of 5th International Conference on Fracture Mechanics of Concrete Structures, Ia-FRAMCOS*, vol. 2, 2004, pp. 949–956
- W. Gerstle, N. Sau, S. Silling, Peridynamic modeling of plain and reinforced concrete structures, in *SMIRT18: 18th International Conference on Structural Mechanics in Reactor Technology*, Beijing, 2005
- Q. Guan, M. Gunzburger, Analysis and approximation of a nonlocal obstacle problem. *J. Comput. Appl. Math.* **313**, 102–118 (2017)
- D. Littlewood, Simulation of dynamic fracture using peridynamics, finite element analysis, and contact, in *Proceeding of ASME 2010 International Mechanical Engineering Congress and Exposition*, Vancouver, 2010
- T. Mengesha, Q. Du, Analysis of a scalar nonlocal peridynamic model with a sign changing kernel. *Disc. Cont. Dyn. Syst. B* **18**, 1415–1437 (2013)
- T. Mengesha, Q. Du, The bond-based peridynamic system with Dirichlet type volume constraint. *Proc. R. Soc. Edinb. A* **144**, 161–186 (2014a)
- T. Mengesha, Q. Du, Nonlocal constrained value problems for a linear peridynamic Navier equation. *J. Elast.* **116**, 27–51 (2014b)
- T. Mengesha, Q. Du, Characterization of function spaces of vector fields via nonlocal derivatives and an application in peridynamics. *Nonlinear Anal. A Theory Meth. Appl.* **140**, 82–111 (2016)
- M. Parks, R. Lehoucq, S. Plimpton, S. Silling, Implementing peridynamics within a molecular dynamics code. *Comput. Phys. Commun.* **179**, 777–783 (2008)
- M. Parks, D. Littlewood, J. Mitchell, S. Silling, *Peridigm Users' Guide*, Sandia report 2012–7800, Albuquerque, 2012
- P. Seleson, M. Parks, M. Gunzburger, R. Lehoucq, Peridynamics as an upscaling of molecular dynamics. *Mult. Model. Simul.* **8**, 204–227 (2009)
- P. Seleson, Q. Du, M. Parks, On the consistency between nearest-neighbor peridynamic discretizations and discretized classical elasticity models. *Comput. Meth. Appl. Mech. Engrg.* **311**, 698–722 (2016)
- S. Silling, Reformulation of elasticity theory for discontinuities and long-range forces. *J. Mech. Phys. Solids* **48**, 175–209 (2000)
- S. Silling, Dynamic fracture modeling with a meshfree peridynamic code, in *Computational Fluid and Solid Mechanics* (Elsevier, Amsterdam, 2003), pp. 641–644
- S. Silling, Linearized theory of peridynamic states. *J. Elast.* **99**, 85–111 (2010)
- S. Silling, E. Askari, Peridynamic modeling of impact damage, in *PVP-Vol. 489* (ASME, New York, 2004), pp. 197–205
- S. Silling, E. Askari, A meshfree method based on the peridynamic model of solid mechanics. *Comput. Struct.* **8**, 1526–1535 (2005)
- S. Silling, F. Bobaru, Peridynamic modeling of membranes and fibers. *Int. J. Nonlinear Mech.* **40**, 395–409 (2005)
- S. Silling, R.B. Lehoucq, Convergence of peridynamics to classical elasticity theory. *J. Elast.* **93**, 13–37 (2008)
- S. Silling, R. Lehoucq, Peridynamic theory of solid mechanics. *Adv. Appl. Mech.* **44**, 73–168 (2010)
- S. Silling, M. Zimmermann, R. Abeyaratne, Deformation of a peridynamic bar. *J. Elast.* **73**, 173–190 (2003)
- S. Silling, M. Epton, O. Weckner, J. Xu, E. Askari, Peridynamic states and constitutive modeling. *J. Elast.* **88**, 151–184 (2007)
- S. Silling, D. Littlewood, P. Seleson, Variable horizon in a peridynamic medium, Technical report No. SAND2014-19088 (Sandia National Laboratories, Albuquerque, 2014)
- X. Tian, Q. Du, Analysis and comparison of different approximations to nonlocal diffusion and linear peridynamic equations. *SIAM J. Numer. Anal.* **51**, 3458–3482 (2013)
- X. Tian, Q. Du, Asymptotically compatible schemes and applications to robust discretization of nonlocal models. *SIAM J. Numer. Anal.* **52**, 1641–1665 (2014)
- X. Tian, Q. Du, Nonconforming discontinuous Galerkin methods for nonlocal variational problems. *SIAM J. Numer. Anal.* **53**(2), 762–781 (2015)

- X. Tian, Q. Du, Trace theorems for some nonlocal energy spaces with heterogeneous localization. *SIAM J. Math. Anal.* **49**(2), 1621–1644 (2017)
- H. Tian, L. Ju, Q. Du, Nonlocal convection-diffusion problems and finite element approximations. *Comput. Meth. Appl. Mech. Engrg.* **289**, 60–78 (2015)
- X. Tian, Q. Du, M. Gunzburger, Asymptotically compatible schemes for the approximation of fractional Laplacian and related nonlocal diffusion problems on bounded domains. *Adv. Comput. Math.* **42**, 1363–1380 (2016)
- H. Tian, L. Ju, Q. Du, A conservative nonlocal convection-diffusion model and asymptotically compatible finite difference discretization. *Comput. Methods Appl. Mech. Engrg.* **320**, 46–67 (2017)
- H. Wang, H. Tian, A fast Galerkin method with efficient matrix assembly and storage for a peridynamic model. *J. Comput. Phys.* **240**, 49–57 (2012)
- O. Weckner, R. Abeyaratne, The effect of long-range forces on the dynamics of a bar. *J. Mech. Phys. Solids* **53**, 705–728 (2005)
- O. Weckner, E. Emmrich, Numerical simulation of the dynamics of a nonlocal, inhomogeneous, infinite bar. *J. Comput. Appl. Mech.* **6**, 311–319 (2005)
- F. Xu, M. Gunzburger, J. Burkardt, Q. Du, A multiscale implementation based on adaptive mesh refinement for the nonlocal peridynamics model in one dimension. *Multiscale Model. Simul.* **14**, 398–429 (2016a)
- F. Xu, M. Gunzburger, J. Burkardt, A multiscale method for nonlocal mechanics and diffusion and for the approximation of discontinuous functions, *Comput. Meth. Appl. Mech. Engrg.* **307**, 117–143 (2016b)
- X. Zhang, M. Gunzburger, L. Ju, Nodal-type collocation methods for hypersingular integral equations and nonlocal diffusion problems. *Comput. Meth. Appl. Mech. Engrg.* **299**, 401–420 (2016a)
- X. Zhang, M. Gunzburger, L. Ju, Quadrature rules for finite element approximations of 1D nonlocal problems. *J. Comput. Phys.* **310**, 213–236 (2016b)
- K. Zhou, Q. Du, Mathematical and numerical analysis of linear peridynamic models with nonlocal boundary conditions. *SIAM J. Numer. Anal.* **48**, 1759–1780 (2010)



Optimization-Based Coupling of Local and Nonlocal Models: Applications to Peridynamics

36

Marta D'Elia, Pavel Bochev, David J. Littlewood, and Mauro Perego

Contents

Introduction	1224
Structure of the Chapter	1224
Local-to-Nonlocal Coupling Methods for Continuum Mechanics	1225
Principles of Optimization-Based Couplings	1226
Well-Posedness	1227
The State Models and Their Properties	1229
Optimization-Based LtN Formulation of Linearized Linear Peridynamic Solid and Classical Elasticity	1231
Discretization of the LtN Formulation	1233
Software	1233
Numerical Tests	1234
Patch Tests	1234
Rectangular Bar with a Crack	1235
Tensile Test Specimen with a Crack	1238
References	1240

Abstract

Nonlocal continuum theories for mechanics can capture strong nonlocal effects due to long-range forces in their governing equations. When these effects cannot be neglected, nonlocal models are more accurate than partial differential

M. D'Elia (✉)

Optimization and Uncertainty Quantification Department Center for Computing Research, Sandia National Laboratories, Albuquerque, NM, USA

e-mail: mdelia@sandia.gov

P. Bochev · D. J. Littlewood · M. Perego

Center for Computing Research, Sandia National Laboratories, Albuquerque, NM, USA

e-mail: pbboche@sandia.gov; djlittl@sandia.gov; mperego@sandia.gov

© This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply 2019

1223

G. Z. Voyiadjis (ed.), *Handbook of Nonlocal Continuum Mechanics for Materials and Structures*, https://doi.org/10.1007/978-3-319-58729-5_31

equations (PDEs); however, the accuracy comes at the price of a prohibitive computational cost, making local-to-nonlocal (LtN) coupling strategies mandatory.

In this chapter, we review the state of the art of LtN methods where the efficiency of PDEs is combined with the accuracy of nonlocal models. Then, we focus on optimization-based coupling strategies that couch the coupling of the models into a control problem where the states are the solutions of the nonlocal and local equations, the objective is to minimize their mismatch on the overlap of the local and nonlocal problem domains, and the virtual controls are the nonlocal volume constraint and the local boundary condition. The strategy is described in the context of nonlocal and local elasticity and illustrated by numerical tests on three-dimensional realistic geometries. Additional numerical tests also prove the consistency of the method via patch tests.

Keywords

Optimization-based coupling methods · Local-nonlocal coupling · Nonlocal elasticity · Classical elasticity · Peridynamics · Domain decomposition · Finite element method · Particle methods

Introduction

Nonlocal continuum theories such as peridynamics (Silling and Lehoucq 2010) and physics-based nonlocal elasticity (Di Paola et al. 2009) can capture strong nonlocal effects due to long-range forces at the mesoscale or microscale. For problems where these effects cannot be neglected, nonlocal models are more accurate than classical partial differential equations (PDEs) that only consider interactions due to contact. However, the improved accuracy of nonlocal models comes at the price of a computational cost that is significantly higher than that of PDEs.

The goal of local-to-nonlocal (LtN) coupling methods is to combine the computational efficiency of PDEs with the accuracy of nonlocal models. LtN couplings are imperative when the size of the computational domain or the extent of the nonlocal interactions is such that the nonlocal solution becomes prohibitively expensive to compute, yet the nonlocal model is required to accurately resolve small-scale features (such as crack tips or dislocations that can affect the global material behavior). In this context, the main challenge of a coupling method is the stable and accurate merging of two fundamentally different mathematical descriptions of the same physical phenomena into a physically consistent coupled formulation.

Structure of the Chapter

This chapter is organized as follows. In section “[Principles of Optimization-Based Couplings](#)” we present an abstract framework of optimization-based coupling (OBC) methods. In section “[The State Models and Their Properties](#)” we introduce the static peridynamics and the local elasticity state models and describe their properties. In section “[Optimization-Based LtN Formulation of Linearized Linear](#)

Peridynamic Solid and Classical Elasticity” we specialize the OBC approach to the state models, and in section **“Discretization of the LtN Formulation”** we describe its fully discrete formulation; here we also review the discretization scheme for static peridynamics. Finally, in section **“Numerical Tests”** we demonstrate the consistency and efficiency of the coupling method through several numerical tests using Sandia’s agile software components toolkit.

Local-to-Nonlocal Coupling Methods for Continuum Mechanics

The promise of improved physical fidelity at a lower computational cost has attracted significant attention to the coupling of nonlocal and local material models in continuum mechanics. The bulk of the existing methods though is based on some form of *blending* of the two material models. This blending can involve the energies of the two models, their force balance equations, or even their material properties. We describe three examples that are representative of these types of couplings.

The extension of the Arlequin method (Ben Dhia and Rateau 2005) to LtN couplings of continuum mechanics models by Han and Lubineau (2012) is an example of an *energy-blending* approach.

Their method splits the domain into a nonlocal subdomain, where the nonlocal effects are pronounced, and an overlapping local subdomain, where such effects are negligible. The intersection of these domains forms a “gluing area” where the energy of the system is defined as a weighted average of the local and nonlocal energies. At the local and nonlocal complements of the gluing area, the energy is defined according to the models operating in these regions. A Lagrange multiplier enforces compatibility of the kinematics of both models.

The formulation in Seleson et al. (2013) provides an example of a *force-blending* coupling approach. This method couples peridynamics and classical elasticity by using a weighted average of the local and nonlocal force balance equations in the overlap or *bridging* domain. Similar to Han and Lubineau (2012), the method uses the “pure” local and nonlocal force equations in the complements of this domain. The resulting hybrid model satisfies Newton’s third law and is consistent for linear fields with no external forces (i.e., the method passes a linear patch test).

Finally, the morphing approach in Lubineau et al. (2012) is an example of an LtN coupling scheme based on blending, or morphing, the *material properties* of the two models. The method consists in the definition of a single model over the entire domain with an equilibrium equation that accounts for both local and nonlocal interactions through a gradual change in the material properties characterizing the two models in a “morphing” region. In this region local and nonlocal properties are suitably weighted under the constraint of energy equivalence in the overlap of the two domains. In Azdoud et al. (2013) the same authors extend this method to anisotropic continua.

These coupling methods share two common features. First, by blending energies, forces, or material models, they effectively introduce a *hybrid* material description combining the properties of both the local and nonlocal models in the overlap regions. Second, they treat the kinematic compatibility between the models, e.g.,

the equality of their displacements over a suitable interface, as a constraint in a way that is reminiscent of classical domain decomposition methods. In the next section, we describe a general OBC strategy that differs fundamentally from the blending approaches discussed above and offers some distinct computational and theoretical advantages.

Principles of Optimization-Based Couplings

In contrast to the blending methods described earlier, an OBC strategy treats the coupling condition as an optimization objective, which is minimized subject to the model equations acting independently in their respective subdomains. In so doing OBC reverses the roles of the coupling conditions and the governing equations and keeps the latter separate.

In particular, the coupling of local and nonlocal models is effected by couching the LtN coupling into an optimization problem. The objective is to minimize the mismatch of the local and nonlocal solutions on the overlap of their subdomains, the constraints are the associated governing equations, and the controls are the virtual nonlocal volume constraint and the local boundary condition. This approach is inspired by nonstandard optimization-based domain decomposition methods for PDEs (Discacciati et al. 2013; Du 2001; Du and Gunzburger 2000; Gervasio et al. 2001; Gunzburger et al. 1999, 2000; Gunzburger and Lee 2000; Kuberry and Lee 2013). It has also been applied to the coupling of discrete atomistic and continuum models in Olson et al. (2014a,b). This strategy brings about valuable theoretical and computational advantages. For instance, the coupled problem passes a patch test by construction, its well-posedness typically follows from the well-posedness of the constraint equations, and its numerical solution only requires the implementation of the optimization strategy as the local and nonlocal solvers for the state equations can be used as black boxes. For this reason, we refer to OBC methods as *nonintrusive* as opposed to the coupling methods described in section “[Local-to-Nonlocal Coupling Methods for Continuum Mechanics](#),” which are *intrusive* in the sense that their implementation requires modification of the basic governing equations for the local and nonlocal models in the overlap region. In what follows, we present an abstract formulation of OBCs.

Let $\mathcal{L}_n : \mathcal{V}_n \rightarrow \mathbb{R}$ be a nonlocal operator that accurately describes the behavior of the material in a bounded body and let $\mathcal{L}_l : \mathcal{V}_l \rightarrow \mathbb{R}$ be a local operator that describes the material well enough where the nonlocal effects are negligible. We recall that the numerical solution of the accurate nonlocal model is computationally expensive, whereas the one of the local model is, in general, affordable. As in the coupling methods described in section “[Local-to-Nonlocal Coupling Methods for Continuum Mechanics](#),” we solve the nonlocal model where the nonlocality affects the global material behavior and the local problem everywhere else; the challenge is to couple those models at the interfaces or overlaps of their domains. As explained above, we tackle this by solving an optimization problem where we minimize the difference between the local and nonlocal solutions at the interfaces tuning their values on the virtual boundaries and volumes induced by the domain decomposition; see Fig. 1a.

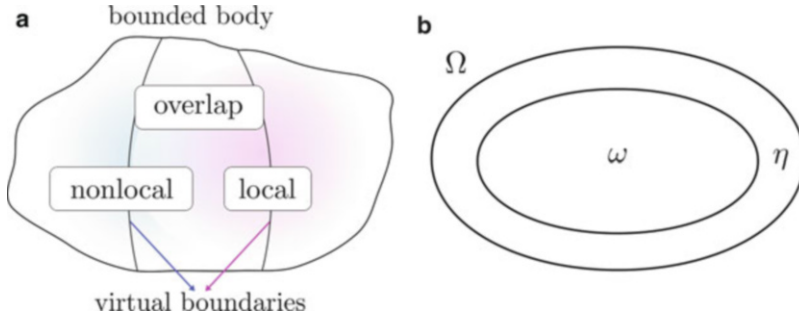


Fig. 1 Illustration of LtN OBC domain configuration for a bonded body Ω and its decomposition into ω and η . (a) An abstract domain configuration. (b) A simplified domain configuration

Formally, we state the LtN OBC as follows:

$$\begin{aligned}
 \min_{\mathbf{u}_n, \mathbf{u}_l, \mathbf{v}_n, \mathbf{v}_l} \mathcal{J}(\mathbf{u}_n, \mathbf{u}_l) &= \frac{1}{2} \|\mathbf{u}_n - \mathbf{u}_l\|_{*, \text{overlap}}^2 \\
 \text{s.t.} \quad \begin{cases} -\mathcal{L}_n \mathbf{u}_n = \mathbf{b}(\mathbf{x}) & \text{nonlocal domain} \\ \mathbf{u}_n(\mathbf{x}) = \mathbf{g}(\mathbf{x}) & \text{physical n-boundary} \\ \mathbf{u}_n(\mathbf{x}) = \mathbf{v}_n(\mathbf{x}) & \text{virtual n-boundary} \end{cases} & \begin{cases} -\mathcal{L}_l \mathbf{u}_l = \mathbf{b}(\mathbf{x}) & \text{local domain} \\ \mathbf{u}_l(\mathbf{x}) = \mathbf{g}(\mathbf{x}) & \text{physical boundary} \\ \mathbf{u}_l(\mathbf{x}) = \mathbf{v}_l(\mathbf{x}) & \text{virtual boundary,} \end{cases} \quad (1)
 \end{aligned}$$

where \mathbf{b} is a body force density, $\|\cdot\|_{*, \text{domain}}$ is a suitable norm on a domain, and $(\mathbf{v}_n, \mathbf{v}_l) \in \mathcal{C}$ (the control space) are the control variables. ‘‘N-boundary’’ stands for nonlocal boundary, usually called interaction volume (rigorously defined in section ‘‘The State Models and Their Properties’’), that consists of all points outside of the domain that interact with points inside the domain. Thus, the goal of OBC is to find optimal values of the virtual controls \mathbf{v}_n and \mathbf{v}_l such that \mathbf{u}_n and \mathbf{u}_l are as close as possible on the overlap and still satisfy the model equations, which play the role of optimization constraints.

Note that this approach is very general and flexible and can be applied to any nonlocal model for continuum mechanics when a suitable local approximation is available. In this chapter, we use the OBC technique to combine nonlocal elasticity, described by a static peridynamics model, and classical linear elasticity. Our strategy is based on the recently introduced approaches (D’Elia and Bochev 2014, 2016; D’Elia et al. 2016) for local and nonlocal diffusion (Du et al. 2012).

Well-Posedness

We present a strategy for proving the well-posedness of (1) for linear operators \mathcal{L}_l and \mathcal{L}_n . Here, without loss of generality, we consider $\mathbf{g} = \mathbf{0}$. We assume that for any pair of controls, the constraints in (1) have unique solutions $\mathbf{u}_n(\mathbf{v}_n)$ and $\mathbf{u}_l(\mathbf{v}_l)$. We introduce the reduced form of the optimization problem by eliminating the states from (1) and obtaining an optimization problem in terms of \mathbf{v}_n and \mathbf{v}_l only:

$$\min_{\mathbf{v}_n, \mathbf{v}_l} \mathcal{J}(\mathbf{v}_n, \mathbf{v}_l) = \frac{1}{2} \|\mathbf{u}_n(\mathbf{v}_n) - \mathbf{u}_l(\mathbf{v}_l)\|_{*,o}^2, \tag{2}$$

where “o” stands for “overlap.” Following the approach used in Abdulle et al. (2015), D’Elia and Bochev (2016), Gervasio et al. (2001) and Olson et al. (2014b), one can show the well-posedness of (2) by splitting the solution of the state equations into the “harmonic” components $(\mathbf{v}_n(\mathbf{v}_n), \mathbf{v}_l(\mathbf{v}_l))$ and the homogeneous components $(\mathbf{u}_n^0, \mathbf{u}_l^0)$ such that they respectively satisfy

$$\left\{ \begin{array}{l} -\mathcal{L}_n \mathbf{v}_n = \mathbf{0} \text{ nonlocal domain} \\ \mathbf{v}_n = \mathbf{0} \text{ physical n-boundary} \\ \mathbf{v}_n = \mathbf{v}_n \text{ virtual n-boundary} \end{array} \right. \quad \left\{ \begin{array}{l} -\mathcal{L}_l \mathbf{v}_l = \mathbf{0} \text{ local domain} \\ \mathbf{v}_l = \mathbf{0} \text{ physical boundary} \\ \mathbf{v}_l = \mathbf{v}_l \text{ virtual boundary,} \end{array} \right. \tag{3}$$

and

$$\left\{ \begin{array}{l} -\mathcal{L}_n \mathbf{u}_n^0 = \mathbf{b} \text{ nonlocal domain} \\ \mathbf{u}_n^0 = \mathbf{0} \text{ n-boundary} \end{array} \right. \quad \left\{ \begin{array}{l} -\mathcal{L}_l \mathbf{u}_l^0 = \mathbf{b} \text{ local domain} \\ \mathbf{u}_l(\mathbf{x}) = \mathbf{0} \text{ boundary.} \end{array} \right. \tag{4}$$

In terms of the components $\mathbf{u}_n = \mathbf{v}_n + \mathbf{u}_n^0$ and $\mathbf{u}_l = \mathbf{v}_l + \mathbf{u}_l^0$, the objective function and the Euler-Lagrange equations are given by

$$\begin{aligned} \mathcal{J}(\mathbf{v}_n, \mathbf{v}_l) &= \frac{1}{2} \|\mathbf{v}_n(\mathbf{v}_n) - \mathbf{v}_l(\mathbf{v}_l)\|_{*,o}^2 + \frac{1}{2} \|\mathbf{u}_n^0 - \mathbf{u}_l^0\|_{*,o}^2 \\ &\quad + (\mathbf{v}_n(\mathbf{v}_n) - \mathbf{v}_l(\mathbf{v}_l), \mathbf{u}_n^0 - \mathbf{u}_l^0)_{*,o}, \end{aligned}$$

and

$$Q(\sigma_n, \sigma_l; \beta_n, \beta_l) = F(\beta_n, \beta_l) \quad \forall (\beta_n, \beta_l) \in \mathcal{C}, \tag{5}$$

where

$$\begin{aligned} Q(\sigma_n, \sigma_l; \beta_n, \beta_l) &= (\mathbf{v}_n(\sigma_n) - \mathbf{v}_l(\sigma_l), \mathbf{v}_n(\beta_n) - \mathbf{v}_l(\beta_l))_{*,o}, \\ F(\beta_n, \beta_l) &= -(\mathbf{u}_n^0 - \mathbf{u}_l^0, \mathbf{v}_n(\beta_n) - \mathbf{v}_l(\beta_l))_{*,o}. \end{aligned}$$

The well-posedness of (2) is a consequence of the following important assumption.

Assumption .1 (Strong Cauchy-Schwarz (CS) inequality) *There exists a positive constant $\kappa < 1$ such that for all $(\sigma_n, \sigma_l) \in \mathcal{C}$*

$$|(\mathbf{v}_n(\sigma_n), \mathbf{v}_l(\sigma_l))_{*,o}| < \kappa \|\mathbf{v}_n(\sigma_n)\|_{*,o} \|\mathbf{v}_l(\sigma_l)\|_{*,o}. \tag{6}$$

This assumption, though strong, is reasonable in the context of multiscale modeling; in fact, it holds for problems such as nonlocal diffusion models (D’Elia and Bochev

2016; Olson et al. 2014b) and multiscale elliptic problems with highly oscillatory coefficients (Abdulle et al. 2015).

The following lemma establishes a fundamental property of Q .

Lemma 1. *If Assumption.1 holds, the form $Q(\cdot, \cdot)$ defines an inner product on \mathcal{C} .*

Proof. The bilinear form $Q(\cdot, \cdot)$ is symmetric and positive semi-definite. We show that it defines an inner product, by showing that $Q(\sigma_n, \sigma_l; \sigma_n, \sigma_l) = 0$ if and only if $(\sigma_n, \sigma_l) = (\mathbf{0}, \mathbf{0})$. Clearly, if $(\sigma_n, \sigma_l) = (\mathbf{0}, \mathbf{0})$, then $\mathbf{v}_n(\sigma_n) = \mathbf{0}$ and $\mathbf{v}_l(\sigma_l) = \mathbf{0}$, implying $Q(\sigma_n, \sigma_l; \sigma_n, \sigma_l) = 0$. On the other hand, if $Q(\sigma_n, \sigma_l; \sigma_n, \sigma_l) = 0$,

$$\begin{aligned} 0 &= Q(\sigma_n, \sigma_l; \sigma_n, \sigma_l) = \|\mathbf{v}_n(\sigma_n) - \mathbf{v}_l(\sigma_l)\|_{*,0}^2 \\ &= \|\mathbf{v}_n(\sigma_n)\|_{*,0}^2 + \|\mathbf{v}_l(\sigma_l)\|_{*,0}^2 - 2(\mathbf{v}_n(\sigma_n), \mathbf{v}_l(\sigma_l))_{*,0} \\ &\geq (1 - \kappa)(\|\mathbf{v}_n(\sigma_n)\|_{*,0}^2 + \|\mathbf{v}_l(\sigma_l)\|_{*,0}^2), \end{aligned}$$

where the last step is a consequence of the strong CS inequality (6) and the Young’s inequality. Since $\kappa < 1$, we have

$$(\|\mathbf{v}_n(\sigma_n)\|_{*,0}^2 + \|\mathbf{v}_l(\sigma_l)\|_{*,0}^2) \leq 0.$$

Thus, we have that $\mathbf{v}_n(\sigma_n) = \mathbf{0}$ and $\mathbf{v}_l(\sigma_l) = \mathbf{0}$, which implies $(\sigma_n, \sigma_l) = (\mathbf{0}, \mathbf{0})$. \square

Note that to establish the well-posedness of problem (2), we need the completeness of \mathcal{C} with respect to the norm induced by Q . However, this may not be the case; thus, as done in Abdulle et al. (2015), we may consider the completion of \mathcal{C} and solve the optimization problem in the completed space, which we denote by \mathcal{C}_c . Then, we use the Hahn-Banach theorem to extend Q and F in \mathcal{C}_c in a continuous and unique way and we denote the extensions by Q_c and F_c . The latter are such that Q_c is continuous and coercive and F_c is continuous in \mathcal{C}_c . The following theorem is a consequence of the considerations above.

Theorem 1. *If Assumption.1 holds, the optimization problem (2) has a unique solution $(\mathbf{v}_n^*, \mathbf{v}_l^*) \in \mathcal{C}_c$ satisfying the extended Euler-Lagrange equation*

$$Q_c(\mathbf{v}_n^*, \mathbf{v}_l^*; \beta_n, \beta_l) = F_c(\beta_n, \beta_l) \quad \forall (\beta_n, \beta_l) \in \mathcal{C}_c.$$

The State Models and Their Properties

Let $\Omega \subset \mathbb{R}^3$ be a bounded body with boundary $\partial\Omega = \Gamma$, the peridynamic equation of the displacement of a material point $\mathbf{x} \in \Omega$ at time $t \geq 0$ is given by

$$\rho(\mathbf{x}, t) \frac{\partial^2 \mathbf{u}}{\partial t^2}(\mathbf{x}, t) = \int_{\Omega} \{ \mathbf{T}[\mathbf{x}, t] \langle \mathbf{x}' - \mathbf{x} \rangle - \mathbf{T}[\mathbf{x}', t] \langle \mathbf{x} - \mathbf{x}' \rangle \} dV_{\mathbf{x}'} + \mathbf{b}(\mathbf{x}, t),$$

where $\rho: \Omega \times \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is the mass density, $\mathbf{u}: \Omega \times \mathbb{R}^+ \rightarrow \mathbb{R}^3$ is the displacement field, $\mathbf{b}: \Omega \times \mathbb{R}^+ \rightarrow \mathbb{R}^3$ is a given body force density and $\mathbf{T}: \Omega \times \mathbb{R}^+ \rightarrow \mathbb{R}^{(3,3)}$ is the force state field, i.e., the force state at (\mathbf{x}, t) mapping the bond $\langle \mathbf{x}' - \mathbf{x} \rangle$ to force per unit volume squared. In this work we consider the peridynamic equilibrium equation for a static problem:

$$-\mathcal{L}_n[\mathbf{u}](\mathbf{x}) := -\int_{\Omega} \{ \mathbf{T}[\mathbf{x}]\langle \mathbf{x}' - \mathbf{x} \rangle - \mathbf{T}[\mathbf{x}']\langle \mathbf{x} - \mathbf{x}' \rangle \} dV_{\mathbf{x}'} = \mathbf{b}(\mathbf{x}). \tag{7}$$

According to the nonlocal theory, we make the assumption that a material point \mathbf{x} interacts only with a neighborhood of points, more specifically, with material points in a ball of radius δ centered in \mathbf{x} , i.e.,

$$B_{\delta}(\mathbf{x}) = \{ \mathbf{x}' \in \Omega : |\mathbf{x} - \mathbf{x}'| \leq \delta \},$$

where δ is a length scale referred to as the *horizon*. This implies that

$$\mathbf{T}[\mathbf{x}]\langle \mathbf{x}' - \mathbf{x} \rangle = 0, \quad \forall \mathbf{x}' \notin B_{\delta}(\mathbf{x}).$$

We solve (7) in $\omega \in \Omega$ and we prescribe Dirichlet volume constraints in a volumetric layer η surrounding ω so that the entire problem domain is $\Omega = \omega \cup \eta$; see Fig. 1, right. The definition of η depends on the properties of \mathbf{T} , and its thickness has to be large enough to guarantee the well-posedness of the problem; we provide more details below. In this work, for simplicity, we consider the linearized linear peridynamic solid (LPS) model (Silling et al. 2007) characterized by the force state field

$$\mathbf{T}[\mathbf{x}]\langle \boldsymbol{\xi} \rangle = \frac{w(|\boldsymbol{\xi}|)}{m} \left\{ (3K - 5G) \theta(\mathbf{x}) \boldsymbol{\xi} + 15G \frac{\boldsymbol{\xi} \otimes \boldsymbol{\xi}}{|\boldsymbol{\xi}|^2} (\mathbf{u}(\mathbf{x} + \boldsymbol{\xi}) - \mathbf{u}(\mathbf{x})) \right\}, \quad \forall \mathbf{x} \in \Omega, \tag{8}$$

where $\boldsymbol{\xi} = \mathbf{x}' - \mathbf{x}$. Here K is the bulk modulus and G is the shear modulus. The linearized nonlocal dilatation, $\theta: \Omega \rightarrow \mathbb{R}$, is defined as

$$\theta(\mathbf{x}) = \frac{3}{m} \int_{B_{\delta}(\mathbf{0})} w(|\boldsymbol{\zeta}|) \boldsymbol{\zeta} (\mathbf{u}(\mathbf{x} + \boldsymbol{\zeta}) - \mathbf{u}(\mathbf{x})) dV_{\boldsymbol{\zeta}}, \quad \text{with } m = \int_{B_{\delta}(\mathbf{0})} w(|\boldsymbol{\zeta}|) |\boldsymbol{\zeta}|^2 dV_{\boldsymbol{\zeta}}.$$

Here, the spherical influence function w is a scalar valued function used to determine the support of force states and to modulate the bond strength (Seleson and Parks 2011; Silling et al. 2007). Using the linearized LPS force state field in (8), we formulate the three-dimensional peridynamic problem as follows. Find $\mathbf{u} \in [L^2(\Omega)]^3$ such that

$$\begin{cases} -\mathcal{L}_{\text{LPS}}[\mathbf{u}](\mathbf{x}) = \mathbf{b}(\mathbf{x}) & \mathbf{x} \in \omega \\ \mathbf{u}(\mathbf{x}) = \mathbf{g}(\mathbf{x}) & \mathbf{x} \in \eta, \end{cases} \tag{9}$$

where $\mathbf{g} \in [L^2(\eta)]^3$ is a given displacement function and \mathcal{L}_{LPS} is obtained by substituting (8) into \mathcal{L}_n , i.e.,

$$\begin{aligned} \mathcal{L}_{\text{LPS}}[\mathbf{u}](\mathbf{x}) := & \int_{B_\delta(\mathbf{0})} \frac{w(|\boldsymbol{\xi}|)}{m} \left\{ (3K - 5G)(\theta(\mathbf{x}) + \theta(\mathbf{x} + \boldsymbol{\xi}))\boldsymbol{\xi} \right. \\ & \left. + 30G \frac{\boldsymbol{\xi} \otimes \boldsymbol{\xi}}{|\boldsymbol{\xi}|^2} (\mathbf{u}(\mathbf{x} + \boldsymbol{\xi}) - \mathbf{u}(\mathbf{x})) \right\} dV_{\boldsymbol{\xi}}. \end{aligned} \quad (10)$$

We define the layer η as

$$\eta = \{\mathbf{x}' \in \Omega : |\mathbf{x}' - \mathbf{x}| < 2\delta\} \quad \forall \mathbf{x} \in \Gamma. \quad (11)$$

Note that the thickness is double the size of the horizon; this happens because in order to evaluate the peridynamic operator on a boundary point $\mathbf{x} \in \partial\omega$, we need to evaluate a double integral over $B_\delta(\mathbf{0}) \times B_\delta(\mathbf{0})$, i.e., we need values of the displacement in $B_{2\delta}(\mathbf{x})$.

The model (10) has two important features. First, its local limit (i.e., the limit for $\delta \rightarrow 0$ that corresponds to vanishing nonlocal interactions) is the classical Navier-Cauchy equation (NCE) of static elasticity (Seleson and Littlewood 2016):

$$-\mathcal{L}_{\text{NC}}[\mathbf{u}](\mathbf{x}) := - \left[\left(K + \frac{1}{3}G \right) \nabla(\nabla \cdot \mathbf{u})(\mathbf{x}) + G \nabla^2 \mathbf{u}(\mathbf{x}) \right] = \mathbf{b}(\mathbf{x}), \quad (12)$$

where K , G , and \mathbf{b} are defined as in (8). The latter is equivalent to the linear elasticity equation in terms of the Lamé constants (λ, μ) :

$$\begin{aligned} -\nabla \cdot \boldsymbol{\sigma}[\mathbf{u}](\mathbf{x}) &= \mathbf{b}(\mathbf{x}), \quad \text{where} \\ \boldsymbol{\sigma}[\mathbf{u}] &= \lambda(\nabla \cdot \mathbf{u})\mathbf{I} + \mu(\nabla \mathbf{u} + \nabla \mathbf{u}^T) \\ (\lambda, \mu) &= \left(K - \frac{2G}{3}, G \right), \end{aligned} \quad (13)$$

where \mathbf{I} is the identity tensor. This property suggests that the NC model can approximate fairly well the nonlocal model for sufficiently regular solutions; for this reason, it is the local model of choice in our coupling strategy.

Second, for a quadratic displacement field, the linearized LPS reduces to the classical NCE (see Proposition 1 in Seleson and Littlewood 2016). This property allows us to perform a quadratic patch test; see section “Patch Tests.”

Optimization-Based LtN Formulation of Linearized Linear Peridynamic Solid and Classical Elasticity

Given a domain Ω representing a bounded body, we introduce a partition into a nonlocal subdomain Ω_n and a local subdomain Ω_l , with boundary Γ_l , such that $\Omega_n = \omega_n \cup \eta_n$ and $\Omega_n \cap \Omega_l = \Omega_o \neq \emptyset$; see Fig. 2 for a two-dimensional illustration.

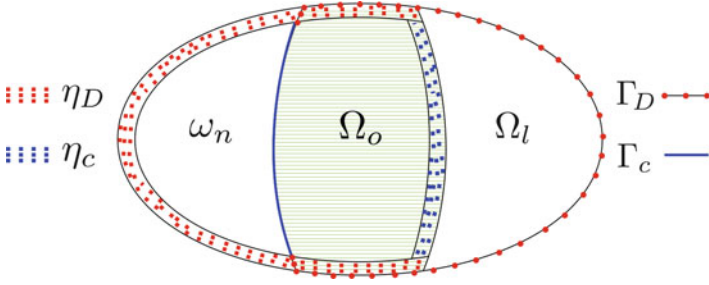


Fig. 2 An example LtN domain configuration in two dimensions

We assume that the nonlocal model (10) accurately describes the material behavior in Ω_n , while the local NC model gives a fairly reasonable representation for the rest of the domain. We formulate the coupling as an optimization problem where we minimize the difference between the nonlocal and the local solutions on the overlap Ω_o adjusting their values on the virtual interaction volume η_c and the virtual boundary Γ_c determined by the partition. Let $\eta_D = \eta \cap \eta_n$ and $\Gamma_D = \Gamma \cap \Gamma_l$ be the physical interaction volume and boundary where we prescribe the given Dirichlet data; we define the virtual control volume and boundary as $\eta_c = \eta_n \setminus \eta_D$ and $\Gamma_c = \Gamma_l \setminus \Gamma_D$. By posing the peridynamic problem on ω_n and the NC problem on Ω_l , we obtain

$$\left\{ \begin{array}{l} -\mathcal{L}_{\text{LPS}}[\mathbf{u}_n](\mathbf{x}) = \mathbf{b}(\mathbf{x}) \quad \mathbf{x} \in \omega_n \\ \mathbf{u}_n(\mathbf{x}) = \mathbf{g}(\mathbf{x}) \quad \mathbf{x} \in \bar{\eta}_D \\ \mathbf{u}_n(\mathbf{x}) = \mathbf{v}_n(\mathbf{x}) \quad \mathbf{x} \in \eta_c \end{array} \right. \left\{ \begin{array}{l} -\mathcal{L}_{\text{NC}}[\mathbf{u}_l](\mathbf{x}) = \mathbf{b}(\mathbf{x}) \quad \mathbf{x} \in \Omega_l \\ \mathbf{u}_l(\mathbf{x}) = \mathbf{g}(\mathbf{x}) \quad \mathbf{x} \in \Gamma_D \\ \mathbf{u}_l(\mathbf{x}) = \mathbf{v}_l(\mathbf{x}) \quad \mathbf{x} \in \Gamma_c, \end{array} \right. \quad (14)$$

where $\mathbf{v}_n \in [L^2(\eta_c)]^3$ and $\mathbf{v}_l \in [H^{1/2}(\Gamma_c)]^3$ are undetermined volume constraints and boundary conditions. In our formulation (14) serve as constraints and $(\mathbf{v}_n, \mathbf{v}_l)$ as control variables of the optimization problem

$$\min_{\mathbf{u}_n, \mathbf{u}_l, \mathbf{v}_n, \mathbf{v}_l} \mathcal{J}(\mathbf{u}_n, \mathbf{u}_l) = \frac{1}{2} \int_{\Omega_o} |\mathbf{u}_n - \mathbf{u}_l|^2 dx \quad \text{subject to (14)}. \quad (15)$$

Given the optimal controls \mathbf{v}_n^* and \mathbf{v}_l^* , we define the coupled solution as

$$\mathbf{u}^* = \begin{cases} \mathbf{u}_n^* & \mathbf{x} \in \Omega_n \\ \mathbf{u}_l^* & \mathbf{x} \in \Omega_l \setminus \Omega_o, \end{cases} \quad (16)$$

where $\mathbf{u}_n^* = \mathbf{u}_n(\mathbf{v}_n^*)$ and $\mathbf{u}_l^* = \mathbf{u}_l(\mathbf{v}_l^*)$.

Discretization of the LtN Formulation

For the discretization of the NC model in (12), we consider the standard finite element (FE) method. We denote the vector of values of the local discrete solution at the FE degrees of freedom by $\vec{u}_l = [\vec{u}_l^1, \vec{u}_l^2, \vec{u}_l^3]$, with $\vec{u}_l^k \in \mathbb{R}^{N_l}$ where N_l is the number of degrees of freedom of each spatial component over the FE computational mesh.

For the peridynamic model introduced in section “The State Models and Their Properties,” we utilize a meshfree discretization. For every point \mathbf{x}_i discretizing the body Ω , we approximate the integral operator as follows

$$L[\mathbf{x}_i] := \sum_{j \in \mathcal{F}_i} \{ \mathbf{T}[\mathbf{x}_i] \langle \mathbf{x}_j - \mathbf{x}_i \rangle - \mathbf{T}[\mathbf{x}_j] \langle \mathbf{x}_i - \mathbf{x}_j \rangle \} V_j^{(i)}, \tag{17}$$

where \mathbf{x}_i and $V_j^{(i)}$ are quadrature points and weights and \mathcal{F}_i represents the set of all points in Ω interacting with the i th material point. Note that the quadrature point \mathbf{x}_j is chosen to coincide with the reference position of the j th node; the quadrature weight $V_j^{(i)}$ is the volume of the intersection between the neighborhood of \mathbf{x}_j and the neighborhood of \mathbf{x}_i , i.e., $|B_\delta(\mathbf{x}_j) \cap B_\delta(\mathbf{x}_i)|$. For \mathbf{x}_j near the boundary of $B_\delta(\mathbf{x}_i)$, $V_j^{(i)}$ represents a partial volume. Details regarding the computation of $V_j^{(i)}$ can be found in Seleson and Littlewood (2016). We denote the vector of values of the discrete nonlocal solution at the material points by $\vec{u}_n = [\vec{u}_n^1, \vec{u}_n^2, \vec{u}_n^3]$, with $\vec{u}_n^k \in \mathbb{R}^{N_n}$, where N_n is the number of material points.

We let $S_n \in \mathbb{R}^{N_o \times N_n}$ be the matrix that selects the components of \vec{u}_n^k in Ω_o and $S_l \in \mathbb{R}^{N_o \times N_l}$ be the operator that evaluates \vec{u}_l^k at the material points in Ω_o ; we define them as

$$(S_n)_{ij} := \delta_{ij} \text{ and } (S_l)_{ij} := \phi_j(\mathbf{x}_i), \quad \forall \mathbf{x}_i \in \Omega_o,$$

where ϕ_j is the j th FE basis function. We define the discrete functional as

$$J_d(\vec{u}_n, \vec{u}_l) = \frac{1}{2} \sum_{i=1}^{N_o} \sum_{k=1}^3 |(S_n \vec{u}_n^k)_i - (S_l \vec{u}_l^k)_i|^2 \tilde{V}_i, \tag{18}$$

where \tilde{V}_i is the volume associated with the i th material point, properly scaled.

Software

The example simulations are carried out using the Albany Salinger et al. (2013) (available at the public git repository <https://github.com/gahansen/Albany>) and Peridigm Parks et al. (2012) (available at the public git repository <https://github.com/peridigm/peridigm>) codes, developed in the Center for Computing Research

at Sandia National Laboratories. *Albany* is an FE code for simulating a variety of physical processes governed by PDEs. It is applied for the majority of the computation, including FE assembly for the Navier-Cauchy equation, calculation of the functional and its derivative, and solution of the state and adjoint systems. *Peridigm* is a peridynamics code for solid mechanics. A software interface was developed to facilitate the linking of *Peridigm* routines with *Albany*; both *Albany* and *Peridigm* rely on several *Trilinos* packages (available at <https://trilinos.org/packages>), for example, *Epetra* for the management of parallel data structures, *Intrepid2* for FE assembly, and *Ipack* and *AztecOO* for the preconditioning and solution of linear systems. We apply the LBFGS optimization algorithm, as implemented in the *Trilinos* package *ROL* (available at <https://trilinos.org/packages/rol>).

Numerical Tests

In this section, we demonstrate the effectiveness and consistency of our strategy through several numerical examples. We first show that the OBC method passes linear and quadratic patch tests. In these cases, the analytic solutions are available and are in agreement for the nonlocal and local models. We then apply the OBC approach to test cases in which a discontinuity is present in the nonlocal domain. For these simulations, while the nonlocal and local models behave similarly, differences in their solutions are expected in the overlap domain. We model a rectangular bar containing a crack, followed by a tensile test specimen containing a crack. The latter case represents a realistic engineering geometry that fully exercises the OBC approach in three dimensions.

Patch Tests

The patch test simulations demonstrate the effectiveness of the OBC approach on benchmark problems for which the analytic solutions are available. As mentioned previously, it was shown in Seleson and Littlewood (2016) that Eqs. (9) and (12) are equivalent for linear and quadratic displacements. As a result, for this class of problems, it is expected that numerical results obtained using the OBC approach should exhibit an excellent match between the local and nonlocal models in the overlap region, with discretization error being the only source of discrepancy.

We consider a rectangular bar in three dimensions:

$$\begin{aligned}\Omega &= [0.0 \text{ mm}, 100.0 \text{ mm}] \times [-12.5 \text{ mm}, 12.5 \text{ mm}] \times [-12.5 \text{ mm}, 12.5 \text{ mm}], \\ \omega_n \cup \eta_D \cup \eta_c &= [0.0 \text{ mm}, 62.5 \text{ mm}] \times [-12.5 \text{ mm}, 12.5 \text{ mm}] \times [-12.5 \text{ mm}, 12.5 \text{ mm}], \\ \Omega_l &= [37.5 \text{ mm}, 100.0 \text{ mm}] \times [-12.5 \text{ mm}, 12.5 \text{ mm}] \times [-12.5 \text{ mm}, 12.5 \text{ mm}].\end{aligned}$$

Following the configuration illustrated in Fig. 2, the nonlocal domains are constructed such that ω_n is fully encapsulated by $\eta_D \cup \eta_c$. The external layer provided by the domain $\eta_D \cup \eta_c$, in which volume constraints are prescribed, has a thickness equal to twice the horizon (see section “The State Models and Their Properties”).

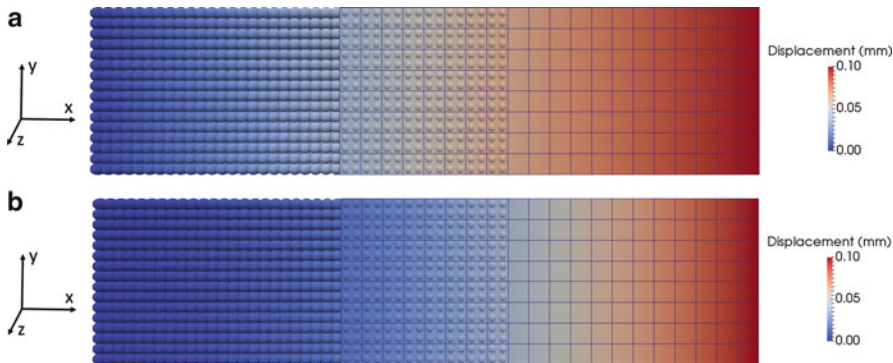


Fig. 3 Solutions for displacement in the x direction for the linear and quadratic patch tests. (a) Solution for the linear patch test. (b) Solution for the quadratic patch test

In Fig. 3, nodal volumes on the left and the FE mesh on the right represent the discretizations of $\omega_n \cup \eta_D \cup \eta_c$ and Ω_l , respectively. Further, we define

$$\begin{aligned} \text{Linear: } \mathbf{u}(\mathbf{x}) &= 10^{-3}(x, 0, 0), \mathbf{b}(\mathbf{x}) = \mathbf{0}, \mathbf{g}(\mathbf{x}) = \mathbf{u}(\mathbf{x}), \\ \text{Quadratic: } \mathbf{u}(\mathbf{x}) &= 10^{-5}(x^2, 0, 0), \mathbf{b}(\mathbf{x}) = \mathbf{b}_q, \mathbf{g}(\mathbf{x}) = \mathbf{u}(\mathbf{x}). \end{aligned}$$

We assign to the bulk modulus, K , a value of 150.0 GPa, and we assign to the shear modulus, G , a value of 81.496 GPa, which are representative of stainless steel. The peridynamic horizon in the nonlocal domain is assigned a value of 4.270 mm. Following Seleson and Littlewood (2016), the body force density, \mathbf{b}_q , producing equilibrium under the given quadratic displacement field, is given by

$$\mathbf{b}_q = 10^{-5} \left(\frac{8G}{3} + 2K \right) = 5.173 \text{ N mm}^{-3}.$$

Simulation results for the linear and quadratic patch tests are presented in Figs. 3 and 4. Displacement solutions in the x (horizontal) direction are given in Fig. 3. In Fig. 4, we report the same variable along a horizontal line passing through the center of the bar. The patch test results are in good agreement with the expected linear and quadratic solutions, respectively, for both the nonlocal and local models.

Rectangular Bar with a Crack

We next consider a rectangular bar containing a discontinuity (crack) at its center. As illustrated in Fig. 5, OBC is utilized to connect a nonlocal domain covering the center portion of the bar with two local domains located at the ends of the bar. Under this configuration, the discontinuity is contained within the nonlocal domain, and the regions over which (non-control) Dirichlet boundary conditions are applied are restricted to the local domain. This is advantageous because, in

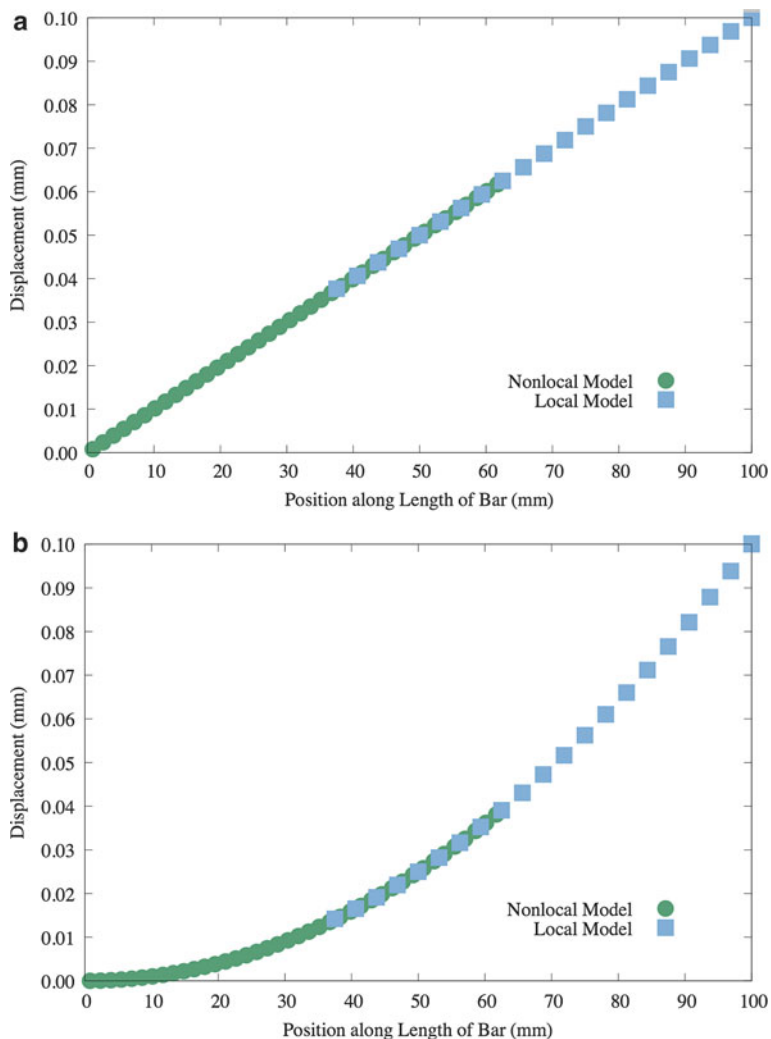


Fig. 4 Solutions for the x component of displacement along a horizontal line passing through the center of the bar for the linear and quadratic patch tests. (a) Solution for the linear patch test. (b) Solution for the quadratic patch test

practice, the determination and application of nonlocal volume constraints can be problematic (Littlewood 2015). We define the bounded body as

$$\Omega := [-50.0 \text{ mm}, 50.0 \text{ mm}] \times [-12.5 \text{ mm}, 12.5 \text{ mm}] \times [-12.5 \text{ mm}, 12.5 \text{ mm}].$$

The discontinuity is inserted into the geometry via a rectangular plane defined by $x = 0.0 \text{ mm}$, $5.0 \text{ mm} \leq y \leq 12.5 \text{ mm}$, and $-12.5 \text{ mm} \leq z \leq 12.5 \text{ mm}$. The nonlocal and local domains are defined as

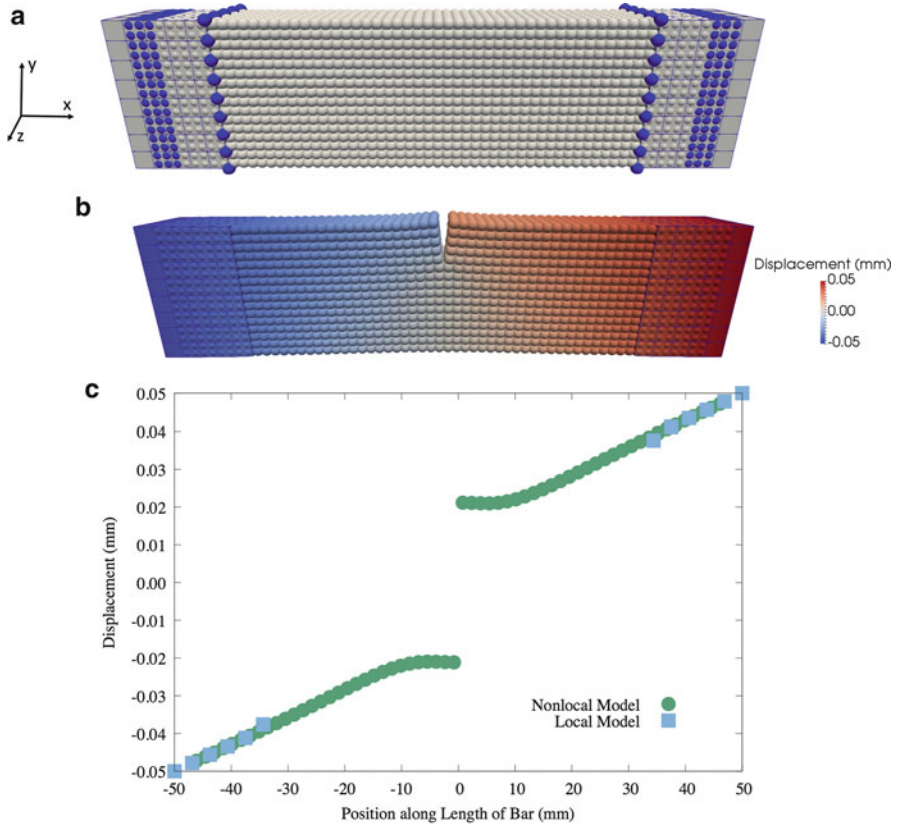


Fig. 5 Discretization and solution for the rectangular bar with a crack. (a) Discretization for the rectangular bar with a crack. Control nodes are highlighted in blue. (b) The x component of the displacement solution. Deformation is magnified by a factor of 20 to clearly illustrate the discontinuity. (c) The x component of the displacement solution along a horizontal line on the top edge of the bar, passing through the discontinuity

$$\begin{aligned} \omega_n \cup \eta_D \cup \eta_c &:= [-46.875 \text{ mm}, 46.875 \text{ mm}] \times [-12.5 \text{ mm}, 12.5 \text{ mm}] \times [-12.5 \text{ mm}, \\ &\quad 12.5 \text{ mm}], \\ \Omega_{l1} &:= [-50.0 \text{ mm}, -34.375 \text{ mm}] \times [-12.5 \text{ mm}, 12.5 \text{ mm}] \times [-12.5 \text{ mm}, 12.5 \text{ mm}], \\ \Omega_{l2} &:= [34.375 \text{ mm}, 50.0 \text{ mm}] \times [-12.5 \text{ mm}, 12.5 \text{ mm}] \times [-12.5 \text{ mm}, 12.5 \text{ mm}]. \end{aligned}$$

The domain η_c , over which control Dirichlet conditions for the nonlocal domain are applied, is defined by $-46.875 \text{ mm} \leq x \leq -42.1875 \text{ mm}$ and $42.1875 \text{ mm} \leq x \leq 46.875 \text{ mm}$. The control Dirichlet conditions for the local model are applied to Γ_c , defined by the planes $x = -34.375 \text{ mm}$ and $x = 34.375 \text{ mm}$. The locations of the control nodes in the discretized model are highlighted in Fig. 5a. As in the patch tests, the bulk modulus, K , is assigned a value of 150.000 GPa and the shear modulus, G , a value of 81.496 GPa. The peridynamic horizon in the nonlocal

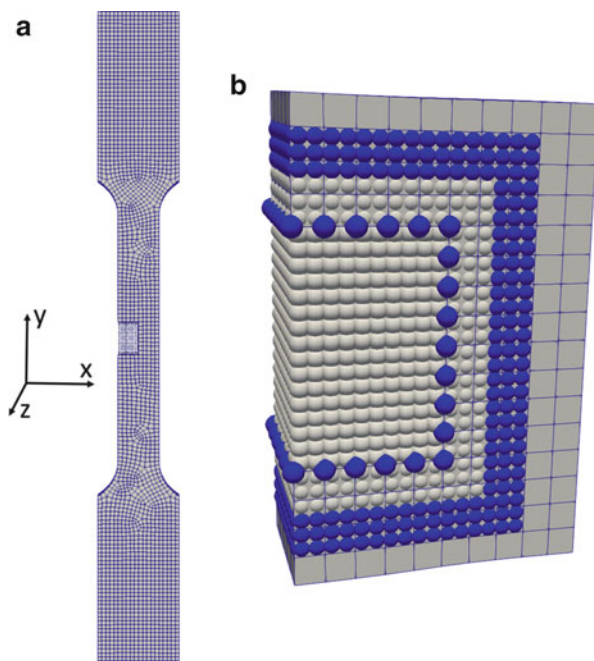
domain is assigned a value of 2.707 mm. Tensile loading is applied to the bar by prescribing displacements of -0.05 mm and 0.05 mm in the x (longitudinal) direction on the faces located at the ends of the bar defined by $x = -50.0$ mm and $x = 50.0$ mm, respectively. To eliminate rigid body modes, additional zero displacement boundary conditions are applied in the y direction along the edges defined by $x = -50.0$ mm, $y = -12.5$ mm and $x = 50.0$ mm, $y = -12.5$ mm and in the z direction along the edges defined by $x = -50.0$ mm, $z = -12.5$ mm and $x = 50.0$ mm, $z = -12.5$ mm.

Simulation results are presented in Fig. 5. The three-dimensional image in Fig. 5b shows the opening of the crack that results from tensile loading. Figure 5c gives displacement results along a horizontal line located on the top face of the bar.

Tensile Test Specimen with a Crack

The simulation of a tensile bar with a crack at its midpoint demonstrates OBC for the modeling of a common engineering geometry. As shown in Fig. 6, we restrict the use of the nonlocal model to a small subdomain in the direct vicinity of the crack. The overall height of the tensile bar specimen is 100.0 mm and the width of the bar at its midpoint is 6.25 mm. The nonlocal region, located at the midpoint of the bar and offset to the side of the bar containing the crack, has a height of 8.68 mm and width of 4.985 mm. The nodes comprising the nonlocal model control domain, η_c ,

Fig. 6 Discretization of the tensile bar specimen. The nonlocal domain is restricted to a small subregion near the center of the bar.
(a) Discretization of tensile bar specimen. (b) Control nodes in the overlap region



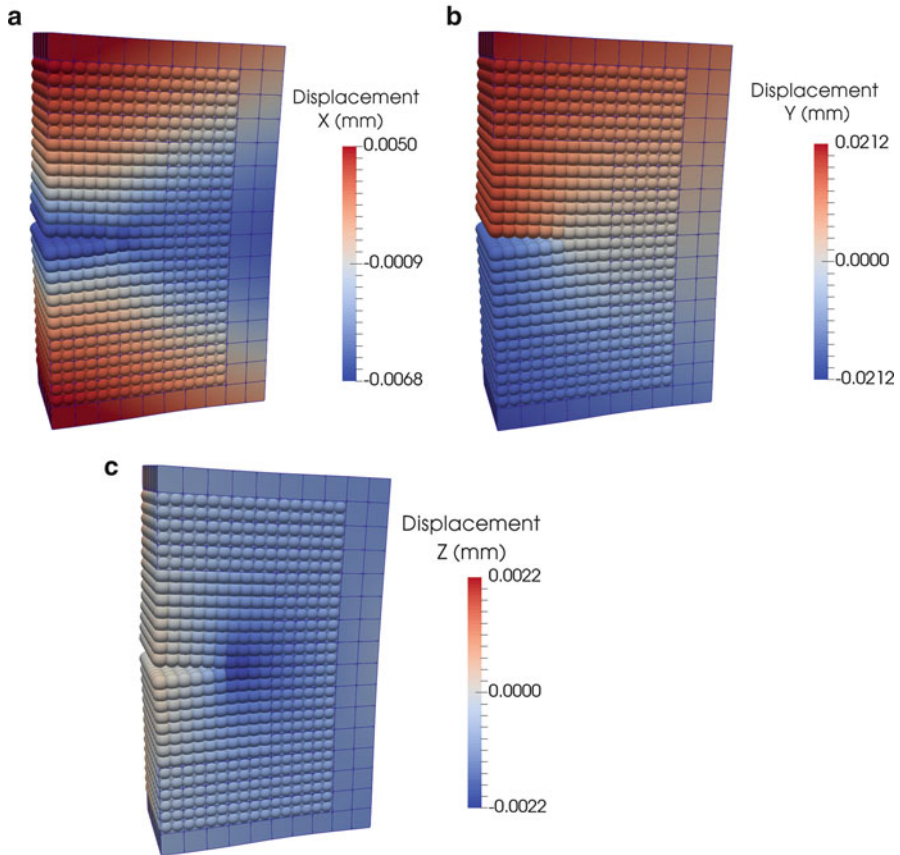


Fig. 7 Displacement solutions for the tensile test specimen. Deformation is magnified by a factor of 10 to clearly illustrate the discontinuity. (a) Displacement in x direction. (b) Displacement in y direction. (c) Displacement in z direction

and the local model control domain, Γ_c , are highlighted in blue in Fig. 6b. The discontinuity is inserted via a rectangular plane at the midpoint of the bar extending from the left side of the bar a distance of 1.86 mm into the bar. We employ material model parameters of 160.0 GPa for the bulk modulus and 64.0 GPa for the shear modulus. For the nonlocal model, the peridynamic horizon is assigned a value of 0.537 mm. Tensile loading is simulated via Dirichlet (displacement) boundary condition applied to the faces at the top and bottom of the bar that produce an overall engineering strain of 0.1% in the y direction. Following the strategy described in section “[Rectangular Bar with a Crack](#),” additional zero displacement boundary conditions are applied along edges on the top and bottom faces in the x and z directions to eliminate rigid body modes.

Results for the tensile bar simulation are given in Fig. 7. The influence of the crack on the displacement solution is restricted predominantly to the nonlocal

region, and solutions corresponding to the nonlocal and local models are in good general agreement in the overlap domain.

Acknowledgements This material is based upon work supported by the US DOE's Laboratory Directed Research and Development (LDRD) program at Sandia National Laboratories and the US Department of Energy, Office of Science, Office of Advanced Scientific Computing Research. Part of this research was carried under the auspices of the Collaboratory on Mathematics for Mesoscopic Modeling of Materials (CM4). Sandia National Laboratories is a multitechnology laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the US Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525. SAND2017-3003 B.

References

- A. Abdulle, O. Jecker, A. Shapeev, An optimization based coupling method for multiscale problems. Technical Report 36.2015, EPFL, Mathematics Institute of Computational Science and Engineering, Lausanne, Dec 2015
- Y. Azdoud, F. Han, G. Lubineau, A morphing framework to couple non-local and local anisotropic continua. *Int. J. Solids Struct.* **50**(9), 1332–1341 (2013)
- M. D'Elia, P. Bochev, Optimization-based coupling of nonlocal and local diffusion models, in *Proceedings of the Fall 2014 Materials Research Society Meeting*, ed. by R. Lipton. MRS Symposium Proceedings (Cambridge University Press, Boston, 2014)
- M. D'Elia, P. Bochev, Formulation, analysis and computation of an optimization-based local-to-nonlocal coupling method. Technical Report SAND2017-1029J, Sandia National Laboratories, 2016
- M. D'Elia, M. Perego, P. Bochev, D. Littlewood, A coupling strategy for nonlocal and local diffusion models with mixed volume constraints and boundary conditions. *Comput. Math. Appl.* **71**(11), 2218–2230 (2016)
- H.B. Dhia, G. Rateau, The Arlequin method as a flexible engineering design tool. *Int. J. Numer. Methods Eng.* **62**(11), 1442–1462 (2005)
- M. Di Paola, G. Failla, M. Zingales, Physically-based approach to the mechanics of strong non-local linear elasticity theory. *J. Elast.* **97**(2), 103–130 (2009)
- M. Discacciati, P. Gervasio, A. Quarteroni, The interface control domain decomposition (ICDD) method for elliptic problems. *SIAM J. Control. Optim.* **51**(5), 3434–3458 (2013)
- Q. Du, Optimization based nonoverlapping domain decomposition algorithms and their convergence. *SIAM J. Numer. Anal.* **39**(3), 1056–1077 (2001)
- Q. Du, M.D. Gunzburger, A gradient method approach to optimization-based multidisciplinary simulations and nonoverlapping domain decomposition algorithms. *SIAM J. Numer. Anal.* **37**(5), 1513–1541 (2000)
- Q. Du, M. Gunzburger, R. Lehoucq, K. Zhou, Analysis and approximation of nonlocal diffusion problems with volume constraints. *SIAM Rev.* **54**(4), 667–696 (2012)
- P. Gervasio, J.-L. Lions, A. Quarteroni, Heterogeneous coupling by virtual control methods. *Numerische Mathematik* **90**, 241–264 (2001). <https://doi.org/10.1007/s002110100303>
- M.D. Gunzburger, H.K. Lee, An optimization-based domain decomposition method for the Navier-Stokes equations. *SIAM J. Numer. Anal.* **37**(5), 1455–1480 (2000)
- M.D. Gunzburger, J.S. Peterson, H. Kwon, An optimization based domain decomposition method for partial differential equations. *Comput. Math. Appl.* **37**(10), 77–93 (1999)
- M.D. Gunzburger, M. Heinkenschloss, H.K. Lee, Solution of elliptic partial differential equations by an optimization-based domain decomposition method. *Appl. Math. Comput.* **113**(2–3), 111–139 (2000)

- F. Han, G. Lubineau, Coupling of nonlocal and local continuum models by the Arlequin approach. *Int. J. Numer. Methods Eng.* **89**(6), 671–685 (2012)
- P. Kuberry, H. Lee, A decoupling algorithm for fluid-structure interaction problems based on optimization. *Comput. Methods Appl. Mech. Eng.* **267**, 594–605 (2013)
- D.J. Littlewood, Roadmap for peridynamic software implementation. Report SAND2015-9013, Sandia National Laboratories, Albuquerque, 2015
- G. Lubineau, Y. Azdoud, F. Han, C. Rey, A. Askari, A morphing strategy to couple non-local to local continuum mechanics. *J. Mech. Phys. Solids* **60**(6), 1088–1102 (2012)
- D. Olson, P. Bochev, M. Luskin, A. Shapeev, Development of an optimization-based atomistic-to-continuum coupling method, in *Proceedings of LSSC 2013*, ed. by I. Lirkov, S. Margenov, J. Wasniewski. Lecture Notes in Computer Science (Springer, Berlin/Heidelberg, 2014a)
- D. Olson, P. Bochev, M. Luskin, A. Shapeev, An optimization-based atomistic-to-continuum coupling method. *SIAM J. Numer. Anal.* **52**(4), 2183–2204 (2014b)
- M.L. Parks, D.J. Littlewood, J.A. Mitchell, S.A. Silling, Peridigm Users' Guide v1.0.0. SAND Report 2012-7800, Sandia National Laboratories, Albuquerque, 2012
- A.G. Salinger, R.A. Bartlett, Q. Chen, X. Gao, G.A. Hansen, I. Kalashnikova, A. Mota, R.P. Muller, E. Nielsen, J.T. Ostien, R.P. Pawlowski, E.T. Phipps, W. Sun, Albany: a component-based partial differential equation code built on Trilinos. SAND Report 2013-8430J, Sandia National Laboratories, Albuquerque, 2013
- P. Seleson, D.J. Littlewood, Convergence studies in meshfree peridynamic simulations. *Comput. Math. Appl.* **71**(11), 2432–2448 (2016)
- P. Seleson, M.L. Parks, On the role of the influence function in the peridynamic theory. *Int. J. Multiscale Comput. Eng.* **9**, 689–706 (2011)
- P. Seleson, S. Beneddine, S. Prudhomme, A force-based coupling scheme for peridynamics and classical elasticity. *Comput. Mater. Sci.* **66**, 34–49 (2013)
- S.A. Silling, R.B. Lehoucq, Peridynamic theory of solid mechanics, in *Advances in Applied Mechanics*, vol. 44 (Elsevier, San Diego, 2010), pp. 73–168
- S.A. Silling, M. Epton, O. Weckner, J. Xu, E. Askari, Peridynamic states and constitutive modeling. *J. Elast.* **88**, 151–184 (2007)



Bridging Local and Nonlocal Models: Convergence and Regularity

37

Mikil D. Foss and Petronela Radu

Contents

Introduction	1244
Higher-Order Nonlocal Models	1246
Definitions and Setup	1247
Nonlocal and Local Laplacians	1249
Elliptic Properties for the Nonlocal Laplacian	1249
Scaling of the Nonlocal Laplacian and Pointwise Estimates	1250
Regularity of Nonlocal Solutions for Nonlinear Systems	1254
Higher Integrability	1255
Higher Differentiability	1256
Regularity for Higher-Order Nonlocal Problems	1261
References	1262

Abstract

As nonlocal models become more widespread in applications, we focus on their connections with their classical counterparts and also on some theoretical aspects which impact their implementation. In this context we survey recent developments by the authors and prove some new results on regularity of solutions to nonlinear systems in the nonlocal framework. In particular, we focus on semilinear problems and also on higher-order problems with applications in the theory of plate deformations.

The second author acknowledges support from the Simons Foundation.

M. D. Foss (✉) · P. Radu

Department of Mathematics, University of Nebraska-Lincoln, Lincoln, NE, USA

e-mail: mfoss@math.unl.edu; pradu@math.unl.edu

Keywords

Nonlocal operators · Classical differentiability · Higher integrability · Weakly integrable kernels · Peridynamics

Introduction

The need for improved models in continuum mechanics has motivated rapid developments in partial differential equations (PDEs). Over the last two decades, the success of nonlocal models has given a fresh impetus to investigate partial integrodifferential equations (PIDEs), for which efforts are under way toward formulating a counterpart to the local theory in PDEs. Recent results in the nonlocal framework demonstrate the importance of robust implementations for models that depend on a length scale parameter (Tian and Du 2014). Other situations have shown the critical need for a deep understanding of theoretical aspects, such as well-posedness and regularity of solutions, and also continuity with respect to initial data which guarantee that numerical simulations produce physically relevant solutions. Toward this aim, we present below some existing and new results for the nonlocal theory, which give a correspondence between the nonlocal and the classical elliptic theory for the second- and fourth-order levels.

Recently nonlocal theories have successfully modeled singular phenomena (such as fracture), as well as phenomena with nonlocal features (aggregation models in biology (Sun et al. 2012; Mogilner and Edelstein-Keshet 1999), thermal diffusion (Oterkus et al. 2014), image processing (Gilboa and Osher 2008), sandpile formation, and more; see the monograph (Andreu-Vaillo et al. 2010) and the references therein). The theory of peridynamics (Silling 2000) introduced a model for elastic deformations in which an integral operator collects a cumulative response of pairwise interactions within a neighborhood of each material point. This formulation is applicable to both continuous deformations and those with dynamic fracture. In the absence of fracture, for a domain $\Omega \subset \mathbb{R}^n$ with smooth boundary and collar Γ (a closed set surrounding Ω of positive measure, which will be defined more precisely below), the deformation $u : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ satisfies the steady-state system

$$\begin{cases} \mathcal{L}_\mu u(x) = f(x), & x \in \Omega \\ u(x) = 0, & x \in \Gamma. \end{cases} \quad (1)$$

The above operator, \mathcal{L}_μ , is a nonlocal Laplacian, defined as

$$\mathcal{L}_\mu u(x) := \int_{\Omega \cup \Gamma} [u(y) - u(x)] \mu(x, y) \, dy, \quad (2)$$

with $\mu : (\Omega \cup \Gamma) \times (\Omega \cup \Gamma) \rightarrow \mathbb{R}$ a measurable kernel. The source term satisfies $f \in L^2(\Omega)$, so the equalities in (1) and in the sequel are in the almost everywhere sense.

Recent investigations have focused on the case when μ is integrable and has a finite support radius δ , usually referred to as the horizon of interaction. (If μ is not integrable, then the operator \mathcal{L}_μ needs to be defined by using the principal value). An integrable kernel (also called a weakly singular kernel) is often required in numerical implementations in order to avoid quadrature errors, especially when the function u does not have any a priori smoothness. Indeed, in the case of a highly singular kernel μ , a non-smooth or discontinuous function u will produce an unbounded $\mathcal{L}_\mu u$ with unrectifiable errors in numerical computations. Kernels with non-integrable singularity have been considered before (Caffarelli et al. 2014; Foss and Geisbauer 2012); however, their analysis is very different as it can still appeal to compactness and smoothness results that are not available when $\mu \in L^1$. Here we aim to provide a general and versatile roadmap for the study of regularity of solutions to nonlocal systems and establish connections with the local theory by formulating equivalent nonlocal theorems to existing local results. The bridge between the two theories also involves an analysis of convergence of low-regularity nonlocal solutions to the smooth solutions of the classical solutions as the support of the kernel shrinks to zero. We also prove some new regularity results which provide the groundwork for a more comprehensive regularity theory of nonlocal systems. More precisely, we establish regularity of solutions for nonlocal semilinear systems, as well as for higher-order problems.

Nonlocal models provide an *alternative* description for physical situations modeled by functions that lack any weak differentiability. Thus a nonlocal elliptic problem admits L^2 -regular solutions whenever the forcing is in L^2 , whereas the classical Laplacian endows the solutions with H^2 -regularity (basically, existence of two weak derivatives). This fact is seen as a strong advantage to employing nonlocal models when the physics of the problems show that irregular solutions could arise. An immediate implication in the mathematical analysis shows that all estimates could be derived at the L^2 -level, so no additional approximations or smoothing techniques need to be employed. The techniques for investigations of low-regularity solutions are still under development, but sufficiently many existing results indicate that the nonlocal theory mirrors in many aspects the classical PDE framework. The results that connect the nonlocal theory with the local theory are at two levels:

- Nonlocal counterparts to existing theorems in the local theory are derived (e.g., mean value theorems, regularity of “elliptic” solutions, Poincaré inequality, etc.)
- Convergence results that show that nonlocal operators and/or nonlocal solutions converge to their classical counterparts under appropriate assumptions as the horizon of interaction (the support of the kernel μ) shrinks to zero.

The new paradigm requires that one studies nonlocal solutions u_δ (dependent on the kernel’s horizon, δ) with low-regularity tools. One can then investigate the convergence of u_δ to u as the horizon $\delta \rightarrow 0^+$ and identify a connection to classical results. The table below contrasts and compares the two approaches: in the local setting, smooth approximations of the solutions to the PDE will converge in H^2 to

a weak solution $u \in H^2$, whereas in the nonlocal setting, the rough solutions u_δ of the counterpart PIDE will converge in L^2 to the same H^2 regular solution.

Framework	Convergence
Local	Regularized approximations $\xrightarrow{H^2}$ weak H^2 solution
Nonlocal	L^2 approximations $u_\delta \xrightarrow{L^2}$ weak H^2 solution u

The choice of the model (PDE vs. PIDE) is dictated by the physics and whether it is appropriate to require weakly differentiable solutions that satisfy a PDE or to allow solutions with stronger discontinuities as is possible with a PIDE.

Higher-Order Nonlocal Models

In Radu et al. (2017) the authors introduced and analyzed a system involving a nonlocal biharmonic operator with boundary conditions which are counterparts of the classical boundary (hinged and clamped) conditions. This nonlocal model provides an alternative to classical plate theory, which though well-developed for nice domains (convex or of class C^2) and H^2 solutions, has significant shortcomings regarding irregular solutions and domains. The diagram below summarizes the well-posedness results of Radu et al. (2017) for fourth-order nonlocal problems on arbitrary domains, even in the absence of Lipschitz regularity for the boundary of the domain. When the domains have more regularity, these results also establish a clear connection, as the horizon goes to zero, to the classical theory through L^2 strong convergence.

Regularity of domain	Results for the nonlocal biharmonic system
None	Nonlocal problem well-posed in $L^2(\Omega)$
Class C^1 (relaxation may be possible)	$L^2(\Omega)$ -convergence to distributional solution
Class C^2 or convex C^1	Convergence to <i>weak</i> solution of elliptic hinged or clamped problem
Class C^4	Convergence to <i>regular</i> solution of elliptic hinged or clamped problem

For all nonlocal theories, of second- or higher-order, our goal remains to propose and investigate models that are robust in the presence of singularities. The ensuing nonlinear analysis will carry to applications such as suspension bridges where many new models and interesting advances have recently appeared (Berchio et al. 2016; Gazzola 2015; Radu et al. 2014). By introducing higher-order systems in the

nonlocal setting, we are also considering applications in other disciplines where such systems play an essential role, such as nonlocal diffusion (see Andreu-Vaillo et al. 2010 and the references therein). In phase transitions, the classical Cahn-Hilliard equations are fourth-order systems, which have long been recognized to be approximations to phase separations with a binary mixture, yet the model considers it equipped with continuity or even smoothness. Different nonlocal models (Bates and Han 2005a,b) have already been considered as well, yet their convergence to the classical counterpart has not been established yet.

At a theoretical level, we state in this manuscript the first (to our knowledge) regularity results for solutions to nonlinear second-order nonlocal systems and, also, for higher-order nonlocal systems. The classical theory of these systems is very well-developed (e.g., for recent developments, see Mayboroda and Maz'ya 2014). There are several facts that motivate carrying this study in the nonlocal setting. First, as we will see below, the nonlocal elliptic theory contains many parallel results and properties comparable to the classical theory: energy minimization, comparison principles, and mean value properties (Foss and Radu 2016; Hinds and Radu 2012). Also, as discussed above, nonlocal solutions, seen as “rough” approximations of classical solutions, converge to their classical (smooth) counterparts even for higher-order systems. These convergence results have the potential to solve open classical problems where one could study first its nonlocal counterpart and then investigate the limit of the nonlocal solutions as the horizon shrinks. Finally, a major advantage of studying nonlocal problems is the fact that they obviate the need for smoothness assumptions for solutions of systems of *any* order while still capturing the physical phenomenon.

The rest of this paper is organized as follows. In the next section, we present some results regarding nonlocal ellipticity which provide a foundation for carrying out theoretical investigations of nonlocal problems. In section “[Definitions and Setup](#)” we introduce notation and some general assumptions, after which we present a method for obtaining the scaling needed for obtaining pointwise error estimates between nonlocal and local Laplacians applied to a smooth function; the section concludes with a nonlocal version of Poincaré’s inequality. In section “[Regularity of Nonlocal Solutions for Nonlinear Systems](#)” we review some existing regularity results from the linear case, after which we prove some new results for nonlinear problems for second-order nonlocal problems. In the last section, we use an iteration argument to study regularity of solutions for higher-order systems studied in connection with plate deformations.

Definitions and Setup

For the remainder of the paper, the kernel μ will be a nonnegative, rotationally symmetric, and integrable kernel (Throughout the paper, with an abuse of notation, we may write $\mu(x, y) = \mu(|x - y|)$ to indicate that μ depends only on the distance between x and y). At various points in the sequel, we will refer to the following:

Assumptions and definitions for the kernels:

(A1) $\mu : [0, \infty) \rightarrow [0, \infty)$ is measurable, and there exist $\beta \in [0, n)$ and $c, \delta_0 > 0$ such that

$$\begin{aligned} 0 < \mu(r) &\leq cr^{-\beta}, & \text{for all } r \in (0, \delta_0) \\ \mu(r) &= 0, & \text{for all } r \geq \delta_0. \end{aligned}$$

(A2) There is a $c_0 > 0$ such that $\mu(r) > c_0$, for all $r \in (0, \delta_0]$.

(A3) For every $\delta \in (0, \delta_0]$, define $\mu_\delta : [0, \infty) \rightarrow [0, \infty)$ by

$$\mu_\delta(r) := \begin{cases} \mu(r), & 0 < r \leq \delta \\ 0, & r > \delta. \end{cases}$$

The δ parameter is called the horizon of the kernel. The assumption (A1) ensures the integrability of the kernels μ and μ_δ . The nonlocal Poincaré inequality (Theorem 1) requires (A2).

Denote by ω_n the volume of the unit ball in n dimensions, so $n\omega_n$ will be its surface area. The n dimensional ball of radius $\varepsilon > 0$ centered at x will be denoted by $B_\varepsilon(x)$.

Let $\Omega \subset \mathbb{R}^n$ be an open and bounded domain. With δ_0 provided in (A1), the collars of Ω are the sets

$$\Gamma := \overline{\bigcup_{x \in \Omega} B_{\delta_0}(x) \setminus \Omega}$$

and for each $\delta \in (0, \delta_0]$

$$\Gamma_\delta := \overline{\bigcup_{x \in \Omega} B_\delta(x) \setminus \Omega},$$

where $\delta_0 > 0$ is specified in assumption (A1). For each $\varepsilon > 0$, set

$$\Omega_\varepsilon := \{x \in \Omega : B_\varepsilon(x) \subseteq \Omega\}.$$

We use ∇u for the spatial gradient of a function $u : \mathbb{R}^n \rightarrow \mathbb{R}$, provided it exists. For the mean value of u over a set $E \subset \mathbb{R}^n$ with positive and finite measure $|E|$, we will use the standard notation

$$u_E := \int_E u(x) dx = \frac{1}{|E|} \int_E u(x) dx. \tag{3}$$

Given $\phi \in C^\infty(\mathbb{R}^n)$ and $u \in L^2(\Omega \cup \Gamma)$, define the convolution product as

$$(\phi * u)(x) := \int_{\mathbb{R}^n} \phi(y - x) \bar{u}(y) dy$$

where

$$\bar{u}(y) := \begin{cases} u(y), & y \in \Omega \cup \Gamma \\ 0, & y \notin \Omega \cup \Gamma. \end{cases}$$

Nonlocal and Local Laplacians

In this section, we present some results that provide a connection between the local and nonlocal Laplacian operators.

Elliptic Properties for the Nonlocal Laplacian

The classical elliptic theory has developed to a level where systems with complex nonlinearities and space-dependent coefficients are very well understood, even for domains with fairly rough boundaries. Nonlocal operators have been introduced as natural generalizations of classical differential operators, and their successful implementation in applications continues to promote their theoretical analysis to deeper levels. For second-order problems, several results have been proved that show the intimate connection between nonlocal and local frameworks. More specifically, the following elliptic-type properties hold in the nonlocal framework (Du et al. 2012; Foss and Radu 2016; Hinds and Radu 2012):

Proposition 1. *Suppose that μ satisfies assumption (A1). The operator \mathcal{L}_μ admits the following properties:*

- (a) *If $u \equiv \text{constant}$ then $\mathcal{L}_\mu u = 0$ (trivially). Conversely, if $\mathcal{L}_\mu u = 0$ and $u = 0$ on Γ , then $u \equiv \text{constant}$.*
- (b) *Let $x \in \Omega \cup \Gamma$. For any maximal point x_0 that satisfies $u(x_0) \geq u(x)$, we have $-\mathcal{L}_\mu u(x_0) \geq 0$. Similarly, if x_1 is a minimal point such that $u(x_1) \leq u(x)$, then $-\mathcal{L}_\mu u(x_1) \leq 0$.*
- (c) *$-\mathcal{L}_\mu u$ is a positive semidefinite operator, i.e., $\langle -\mathcal{L}_\mu u, u \rangle \geq 0$, where $\langle \cdot, \cdot \rangle$ denotes the $L^2(\Omega \cup \Gamma)$ inner product.*
- (d)
$$\int_{\Omega \cup \Gamma} \mathcal{L}_\mu u(x) dx = 0.$$
- (e) *If the kernel μ is given by a combination of derivatives of the Dirac mass, then for this particular choice of kernel, we have $\mathcal{L}_\mu = \Delta u$ in the sense of distributions (see Du et al. 2012);*
- (f) *As the radius of the horizon δ goes to zero, the nonlocal solutions converge to their classical counterparts. This fact has been proven for the Laplacian (Mengesha and Du 2014), as well as for the biharmonic operator (Radu et al. 2017);*

- (g) *Dirichlet’s principle and variational arguments apply for nonlocal systems as proven in Hinds and Radu (2012);*
- (h) *A weighted mean value theorem is available; indeed, one can easily see that a nonlocal harmonic function u , i.e., a function u for which $\mathcal{L}_\mu u(x) = 0$ for $x \in \Omega$, must also satisfy*

$$u(x) = \frac{\int_{\Omega \cup \Gamma} \mu(|x - y|)u(y) \, dy}{\int_{\Omega \cup \Gamma} \mu(|x - y|) \, dy}, \quad x \in \Omega.$$

Using this result, one can prove

- (i) $\Delta u = 0$ if and only if $\mathcal{L}_\mu u(x) = 0$ for every kernel $\mu(x, y) = \mu(|x - y|) \in L^1$.

We conclude this subsection by stating the following nonlocal Poincaré-type inequality, versions of which can be found in Aksoylu and Parks (2011) or Hinds and Radu (2012):

Theorem 1 (Nonlocal Poincaré inequality). *Let Ω be an open, bounded domain and $\mu \in L^1(\mathbb{R}^n)$ a nonnegative kernel that satisfies (A2) and (A3). If $u \in L^2(\Omega \cup \Gamma)$, then there is a constant $\lambda_p(\delta, n, \Omega) > 0$ s.t.*

$$\lambda_p \int_{\Omega} |u(x)|^2 \, dx \leq \int_{\Omega} \int_{\Omega \cup \Gamma_\delta} \mu_\delta(|x - y|)|u(y) - u(x)|^2 \, dy \, dx + \int_{\Gamma_\delta} |u(x)|^2 \, dx.$$

Scaling of the Nonlocal Laplacian and Pointwise Estimates

In this subsection, we will assume that (A1)–(A3) hold for the kernels μ and μ_δ . We will derive a scaling that normalizes the nonlocal Laplacian. To facilitate identification of the scaling, a modified definition for \mathcal{L}_{μ_δ} will be used and, for convenience, denoted by \mathcal{L}_δ

$$\mathcal{L}_\delta u(x) = \sigma(\delta) \int_{\Omega \cup \Gamma_\delta} [u(y) - u(x)]\mu_\delta(|x - y|) \, dy. \tag{4}$$

Here $\sigma(\delta)$ is the scaling factor in terms of the support of μ_δ . This horizon-dependent scaling σ will be determined later. Moreover, we will show that the scaled $\mathcal{L}_\delta u$ converges to the local Laplacian Δu at the rate δ^2 , i.e.,

$$|\mathcal{L}_\delta u - \Delta u| \leq C(u)\delta^2,$$

where $C(u)$ depends on bounds for the fourth-order derivatives of u . This error bound between the local and nonlocal Laplacian has been derived before (see, e.g.,

Du et al. 2013). Below we outline the argument that was first presented in Foss and Radu (2016) which has already been utilized in deriving error estimates between nonlocal solutions and their local equivalents for systems that involve a nonlocal biharmonic operator (Radu et al. 2017) and for the newly introduced state Laplacian (Radu and Wells 2017). Additionally, the argument accommodates very general kernels (that are integrable) and shows the optimality of the δ^2 convergence that is observed for C^4 functions.

For each $\delta \in (0, \delta_0]$, we define the function $\pi_\delta : (0, \infty) \rightarrow [0, \infty)$

$$\pi_\delta(r) := \int_\delta^r \rho \mu_\delta(\rho) \, d\rho. \tag{5}$$

Thus, for example, if $\mu(r) = r^{-\beta}$ on $(0, \delta_0]$, then we find that

$$\pi_\delta(r) = \begin{cases} \frac{1}{2-\beta}[r^{2-\beta} - \delta^{2-\beta}], & 0 < r \leq \delta, \beta \neq 2 \\ \ln r - \ln \delta, & 0 < r \leq \delta, \beta = 2 \\ 0, & r > \delta. \end{cases}$$

Assume that $u \in C^4(\Omega \cup \Gamma)$. By using the fundamental theorem of calculus, we have that

$$\begin{aligned} \mathcal{L}_\delta u(x) &= \sigma(\delta) \int_{B_\delta(x)} [u(y) - u(x)] \mu_\delta(|x - y|) \, dy \\ &= \sigma(\delta) \int_{B_\delta(x)} \int_0^1 \nabla u(x + s(y - x)) \cdot (y - x) \mu_\delta(|y - x|) \, ds \, dy. \end{aligned}$$

After changing the order of integration in the double integral and changing variables $z = y - x$, we obtain that

$$\mathcal{L}_\delta u(x) = \sigma(\delta) \int_0^1 \int_{B_\delta(0)} \nabla u(x + sz) \cdot [z \mu_\delta(|z|)] \, dz \, ds.$$

With π_δ given by (5), we have that

$$\nabla_z \pi_\delta(|z|) = \pi'_\delta(z) \frac{z}{|z|} = z \mu_\delta(|z|),$$

so

$$\mathcal{L}_\delta u(x) = \sigma(\delta) \int_0^1 \int_{B_\delta(0)} \nabla u(x + sz) \cdot \nabla_z \pi_\delta(|z|) \, dz \, ds.$$

After an integration by parts and by using the fact that $\pi_\delta(|z|) = \pi_\delta(\delta) = 0$ for $z \in \partial B_\delta(0)$, we obtain

$$\mathcal{L}_\delta u(x) = -\sigma(\delta) \int_0^1 \int_{B_\delta(0)} s \Delta u(x + sz) \pi_\delta(|z|) \, dz \, ds$$

which is equivalent to

$$\begin{aligned} \mathcal{L}_\delta u &= -\sigma(\delta) \int_0^1 \int_{B_\delta(0)} s [\Delta u(x + sz) - \Delta u(x)] \pi_\delta(|z|) \, dz \, ds & (6) \\ &\quad - \Delta u(x) \sigma(\delta) \int_0^1 \int_0^\delta \int_{\partial B_\rho(0)} s \pi_\delta(|\rho|) \, d\omega(z) \, d\rho \, ds. \end{aligned}$$

We compute

$$\int_0^1 \int_0^\delta \int_{\partial B_\rho(0)} s \pi_\delta(\rho) \, d\omega(z) \, d\rho \, ds = \frac{1}{2} n \omega_n \int_0^\delta \pi_\delta(\rho) \rho^{n-1} \, d\rho.$$

This results in the scaling

$$\sigma(\delta) = \frac{-2}{n \omega_n \int_0^\delta \pi_\delta(\rho) \rho^{n-1} \, d\rho} \tag{7}$$

will make the coefficient of the Laplacian to be 1 on the RHS of (6) above.

For example, if $\mu(r) = \begin{cases} r^{-\beta}, & 0 < r \leq \delta \\ 0, & \delta < r \end{cases}$ then for $\beta \neq 2$ we have

$$\sigma(\delta) = \frac{2(2 - \beta + n)}{\omega_n \delta^{2-\beta+n}},$$

where for $n = 1$ we take $\omega_0 = 2$. With the choice of scaling given in (7), we write

$$\mathcal{L}_\delta u(x) - \Delta u(x) = -\sigma(\delta) \int_{B_\delta(0)} \int_0^1 s[\Delta u(x + sz) - \Delta u(x)]\pi_\delta(|z|) \, ds \, dz. \tag{8}$$

By employing the fundamental theorem of calculus and simplifying, we obtain

$$\begin{aligned} \mathcal{L}_\delta u(x) - \Delta u(x) &= -\sigma(\delta) \int_{B_\delta(0)} \int_0^1 \frac{(1-s^2)}{2} [\Delta \nabla u(x + sz)] \cdot z \pi_\delta(|z|) \, ds \, dz \\ &= -\sigma(\delta) \int_{B_\delta(0)} \int_0^1 \frac{(1-s^2)}{2} [\Delta \nabla u(x + sz) - \Delta \nabla u(x)] \cdot z \pi_\delta(|z|) \, ds \, dz, \end{aligned}$$

where we used the fact that

$$\int_{B_\delta(0)} z \pi_\delta(|z|) \, dz = 0.$$

Further we have

$$\mathcal{L}_\delta u(x) - \Delta u(x) = -\sigma(\delta) \int_{B_\delta(0)} \int_0^1 \left(\frac{1}{3} - \frac{s}{2} + \frac{s^3}{6} \right) [\Delta \nabla^2 u(x + sz)z] \cdot z \pi_\delta(|z|) \, ds \, dz.$$

With

$$M_4 = \sup_{x \in \Omega \cup \Gamma} |D^4 u(x)|$$

we estimate

$$\begin{aligned} |\mathcal{L}_\delta u(x) - \Delta u(x)| &\leq \sigma(\delta) M_4 \int_{B_\delta(0)} \int_0^1 \left(\frac{1}{3} - \frac{s}{2} + \frac{s^3}{6} \right) |z|^2 |\pi_\delta(|z|)| \, ds \, dz \\ &= \frac{\sigma(\delta) M_4}{8} n \omega_n \int_0^\delta \rho^{n+1} |\pi_\delta(\rho)| \, d\rho. \end{aligned}$$

Finally, we use the fact that $\rho^{n+1} \leq \rho^{n-1} \delta^2$ to obtain

$$|\mathcal{L}_\delta u(x) - \Delta u(x)| \leq \frac{M_4}{4} \delta^2. \tag{9}$$

This estimate shows that the nonlocal Laplacian approaches the classical Laplacian, as the horizon δ goes to zero, at a rate comparable to δ^2 , independently of the dimension. The convergence also demonstrates that the scaling selected in (7) is the correct one.

Moreover, it was shown that besides the above convergence of the operators (applied to sufficiently smooth functions), one has the convergence of the nonlocal solutions to their classical counterparts. More precisely, it was shown in Mengesha and Du (2014) that the solutions u_δ to the problem

$$\begin{cases} \mathcal{L}_\delta u_\delta(x) = f(x), & x \in \Omega \\ u_\delta(x) = 0, & x \in \Gamma_\delta, \end{cases} \quad (10)$$

obtained for each $\delta \in (0, \delta_0]$, converge strongly in L^2 to u , the solution of the classical Laplace equation

$$\begin{cases} \Delta u(x) = f, & x \in \Omega \\ u(x) = 0, & x \in \partial\Omega. \end{cases}$$

Regularity of Nonlocal Solutions for Nonlinear Systems

In this section we provide some integrability and differentiability properties for solutions of

$$\begin{cases} \mathcal{L}_\mu u(x) = f(x, u(x)), & x \in \Omega \\ u(x) = 0, & x \in \Gamma. \end{cases} \quad (11)$$

The linear problem obtained for $x \mapsto f(x, u(x)) = f(x)$ was investigated in Foss and Radu (2016), where it was shown that a solution's integrability properties are the same as possessed by the function f . This result addressed one of the deficiencies of Poincaré's inequality, mainly, the fact that the integrability of a function cannot be improved based on bounds on its nonlocal gradient, as in the classical framework. In the same setting, we also showed in Foss and Radu (2016) that for a nonlocal Laplacian with integrable kernel (Note that if the kernel is highly singular then there are compact embedding results and Poincaré-type inequalities that do increase the integrability by using estimates on the gradient.) and $f \in C^1(\bar{\Omega})$, solutions to (1) satisfy $u \in W^{1,2}(\Omega)$. Below we will offer an extension to this result by showing that a forcing term $f \in W^{1,2}(\Omega) \cap L^\infty(\Omega)$ will yield a solution u with regularity at the same level.

Higher Integrability

We state first a result that guarantees the higher integrability of solutions to the linear problem (we refer to Foss and Radu (2016) for its proof) and then present an extension to this result in the semilinear setting.

Theorem 2 (Higher integrability of solutions to the linear problem). *Assume $\mu \in L^1(\mathbb{R}^n)$ satisfies (A1). If u is a solution of*

$$\begin{cases} \mathcal{L}_\mu u(x) = f(x), & x \in \Omega \\ u(x) = 0, & x \in \Gamma. \end{cases}$$

with a priori regularity $u \in L^p(\Omega)$, $p > 1$ and we assume that $f \in L^r(\Omega)$, $r > p$ then $u \in L^r(\Omega)$.

The proof for this result is similar, but simpler, than its generalization to a nonlinear setting, which we present below.

Theorem 3 (Higher integrability of solutions to semilinear problems). *Assume $\mu \in L^1(\mathbb{R}^n)$ satisfies (A1) and that given $f : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$, the mapping*

$$u \mapsto g_x(u) = u + f(x, u)$$

is invertible on \mathbb{R} with the inverse g_x^{-1} uniformly Lipschitz continuous with respect to $x \in \mathbb{R}^n$. Then any solution $u \in L^p(\Omega)$, $p > 1$ of (11) satisfies $u \in L^\infty(\Omega)$.

Proof. Without loss of generality, we assume that $\|\mu\|_{L^1} = 1$ and note that the uniform Lipschitz continuity of g_x^{-1} implies there is a constant $M < \infty$ such that

$$|g_x^{-1}(u)| \leq M|u|, \quad \text{for all } u \in \mathbb{R}, x \in \Omega.$$

First, extend any solution u of the above equation by zero outside Ω , and with an abuse of notation, denote the extended function by u as well. In light of our assumptions, we write the integrodifferential equation in (1) as

$$\int_{\mathbb{R}^n} u(y)\mu(|y-x|) dy - u(x) \int_{\mathbb{R}^n} \mu(|y-x|) dy = (u * \mu)(x) - u(x) = f(x, u(x)). \quad (12)$$

From the last equality of (12), we obtain that

$$u(x) = g_x^{-1}((u * \mu)(x)) \quad (13)$$

to which we apply Young's inequality for convolutions to obtain

$$\|u * \mu\|_r \leq \|u\|_p \|\mu\|_q,$$

where $\frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1$, where q is chosen appropriately.

Depending on the power β that controls the growth of the kernel μ , as given by assumption (A1), we have the following cases:

Case 1. For $\beta < n \frac{p-1}{p}$ take $q = \frac{p}{p-1}$ to obtain $r = \infty$. By (13) and using the bound on the growth of g_x^{-1} above, we have that $u \in L^\infty(\Omega)$.

Case 2. If $\beta \geq n \frac{p-1}{p}$, then let $q = \frac{n-\varepsilon}{\beta}$ for $\varepsilon > 0$ arbitrarily small. Then, the degree of integrability of $u * \mu$ is up to (but not including) $r_0 = \frac{\bar{p}n}{n + p(\beta - n)}$; note that $r_0 > 1$ since $p > 1$. Again, by (13) and the bound on g_x^{-1} stated above, this integrability is transferred to u so that $u \in L^r$ for all $r < r_0$. We apply again Young's inequality

$$\|u * \mu\|_{r_1} \leq \|u\|_{r_0} \|\mu\|_{q_1},$$

with $\frac{1}{r_1} = \frac{1}{r_0} + \frac{1}{q_1} - 1$. Take $q_1 = \frac{p(n-\varepsilon)}{(n-\varepsilon)(2p-1) - p\beta}$ to obtain $r_1 = \infty$. Using one more time (13) and the growth of g_x^{-1} , we obtain, as in the previous case, that $u \in L^\infty(\Omega)$.

□

Remark 1. The above theorem is not only dimension independent, it is also applicable for unbounded, as well as bounded, domains.

The integrability result in Theorem 2 is used to establish some differentiability properties of the solution u , which are described next.

Higher Differentiability

The first result shows differentiability for solutions of the linear system (1) when the kernel is differentiable. Our final goal is to establish differentiability of solutions when μ is only assumed to be integrable.

Proposition 2. *If $\mu \in C^1(\mathbb{R}^n)$, $f \in C^1(\Omega)$ and $u \in L^1(\Omega \cup \Gamma)$ is a solution to $\mathcal{L}_\mu u = f$; i.e.,*

$$\int_{\Omega \cup \Gamma} (u(y) - u(x))\mu(|x - y|)dy = f(x), \quad \text{for all } x \in \Omega,$$

then $u \in C^1(\Omega)$.

Proof. The theorem follows easily from the equality

$$u(x) = (u * \mu)(x) - f(x)$$

and the smoothing properties of the convolution operator. \square

In the absence of differentiability assumptions for μ , we will need the following lemmas for the proof of the main result.

Lemma 1. *If $\mu \in L^1(\mathbb{R}^n)$ and $u \in C^1(\Omega \cup \Gamma)$ satisfies*

$$\mathcal{L}_\mu u(x) = f(x), \quad \text{for all } x \in \Omega$$

for $f \in C^1(\Omega)$ then $\mathcal{L}_\mu(\nabla u)(x) = \nabla f(x)$ for all $x \in \Omega$.

Proof. The linearity of the \mathcal{L}_μ operator trivially gives the result. \square

Lemma 2. *If $\mu \in L^1(\mathbb{R}^n)$, $f \in L^1(\Omega)$ and $u \in L^1(\Omega \cup \Gamma)$ satisfies*

$$\mathcal{L}_\mu u(x) = f(x), \quad \text{for all } x \in \Omega,$$

then

$$\mathcal{L}_\mu[\phi * u](x) = (\phi * f)(x), \quad \text{for all } x \in \Omega_\varepsilon.$$

Here $\varepsilon > 0$ and $\phi \in C^\infty(\mathbb{R}^n)$ with $\text{supp}(\phi) \subset B_\varepsilon(0)$

Proof. Again, employ the linearity of \mathcal{L}_μ and of the convolution product. Note that for $x \in \Omega \setminus \Omega_\varepsilon$, the convolution $\phi * u$ may incorporate values of u outside of Ω , so there is no assurance that the equation is satisfied outside of Ω_ε . \square

We are now ready to present the regularity result for linear problems, improving the result given by Theorem 4.3 in Foss and Radu (2016).

Theorem 4 (Regularity for solutions of nonlocal equations). *Let $f \in L^\infty(\Omega) \cap W^{1,2}(\Omega)$. Suppose that $\mu \in L^1(\mathbb{R}^n)$ satisfies assumptions (A1), (A2), and (A3). Then, the unique solution u of the system (10)*

$$\begin{cases} \mathcal{L}_\mu u(x) = f(x), & x \in \Omega \\ u(x) = 0, & x \in \Gamma, \end{cases} \quad (14)$$

satisfies

$$u \in L^\infty(\Omega) \cap W^{1,2}(\Omega).$$

Remark 2. The solutions u to (14) clearly depend on μ , but we suppressed this dependence for the clarity of the argument.

Proof. The existence and uniqueness of $u \in L^2(\Omega \cup \Gamma)$ when $f \in L^2(\Omega)$ were proven before (see Hinds and Radu (2012) for a variational argument). Below we will establish uniform estimates in the gradient which will then yield the desired regularity.

Extend u and f to all of \mathbb{R}^n by 0. Let $\{\phi_\varepsilon\}_{\varepsilon>0}$ be a family of mollifiers defined by $\phi_\varepsilon := \frac{1}{\varepsilon^n} \phi\left(\frac{x}{\varepsilon}\right)$, with $\phi \in C^\infty(\mathbb{R}^n)$ and $\text{supp}(\phi) \subset B_1(0)$. Fix $\varepsilon \in \left(0, \frac{\delta_0}{2}\right)$. By Lemma 2,

$$\mathcal{L}_\mu[\phi_\varepsilon * u](x) = (\phi_\varepsilon * f)(x)$$

for each $x \in \Omega_\varepsilon$. Put $u_\varepsilon := \phi_\varepsilon * u$ and $f_\varepsilon := \phi_\varepsilon * f$. Since $u_\varepsilon \in C^\infty(\Omega \cup \Gamma)$, Lemma 1 implies

$$\mathcal{L}_\mu[\nabla u_\varepsilon](x) = \nabla f_\varepsilon(x) \tag{15}$$

for each $x \in \Omega_\varepsilon$. Since $u = 0$ on Γ , we have that $u_\varepsilon = 0$ on

$$\Omega^\varepsilon := \{x \in \mathbb{R}^n \mid \text{dist}(x, \Omega) < \varepsilon\},$$

hence $\nabla u_\varepsilon = 0$ on $\Omega \cup \Gamma \setminus \Omega^\varepsilon$. We then have from (15) above that for all $x \in \Omega_\varepsilon$

$$\mathcal{L}_\mu[\nabla u_\varepsilon](x) = \int_{\Omega^\varepsilon} [\nabla u_\varepsilon(y) - \nabla u_\varepsilon(x)] \mu(|x - y|) dy = \nabla f_\varepsilon(x).$$

We multiply the equation by ∇u_ε and integrate on Ω_ε to obtain

$$\begin{aligned} & \int_{\Omega_\varepsilon} \int_{\Omega^\varepsilon} [(\nabla u_\varepsilon(y) - \nabla u_\varepsilon(x)) \mu(|y - x|)] \cdot \nabla u_\varepsilon(x) \, dy \, dx \\ &= \int_{\Omega_\varepsilon} \nabla f_\varepsilon(x) \cdot \nabla u_\varepsilon(x) \, dx. \end{aligned} \tag{16}$$

We estimate the LHS above as follows:

$$\begin{aligned} & \int_{\Omega_\varepsilon} \int_{\Omega_\varepsilon} [(\nabla u_\varepsilon(y) - \nabla u_\varepsilon(x)) \mu(|y - x|)] \cdot \nabla u_\varepsilon(x) \, dy \, dx \\ &+ \int_{\Omega_\varepsilon} \int_{\Omega^\varepsilon \setminus \Omega_\varepsilon} [(\nabla u_\varepsilon(y) - \nabla u_\varepsilon(x)) \mu(|y - x|)] \cdot \nabla u_\varepsilon(x) \, dy \, dx \\ &= \frac{1}{2} \int_{\Omega_\varepsilon} \int_{\Omega_\varepsilon} [(\nabla u_\varepsilon(y) - \nabla u_\varepsilon(x))^2 \mu(|y - x|)] \, dy \, dx \end{aligned}$$

$$+ \int_{\Omega_\varepsilon} \int_{\Omega^\varepsilon \setminus \Omega_\varepsilon} [(\nabla u_\varepsilon(y) - \nabla u_\varepsilon(x))\mu(|y-x|)] \cdot \nabla u_\varepsilon(x) \, dy \, dx$$

From (16) and above equality, we obtain

$$\begin{aligned} \int_{\Omega_\varepsilon} \int_{\Omega_\varepsilon} [(\nabla u_\varepsilon(y) - \nabla u_\varepsilon(x))^2 \mu(|y-x|)] \, dy \, dx &= 2 \int_{\Omega_\varepsilon} \nabla f_\varepsilon(x) \cdot \nabla u_\varepsilon(x) \, dx \quad (17) \\ &- 2 \int_{\Omega_\varepsilon} \int_{\Omega^\varepsilon \setminus \Omega_\varepsilon} [(\nabla u_\varepsilon(y) - \nabla u_\varepsilon(x))\mu(|y-x|)] \cdot \nabla u_\varepsilon(x) \, dy \, dx \end{aligned}$$

By Poincaré's inequality as given by Theorem 1 applied on the domain Ω_ε with collar $\Omega^\varepsilon \setminus \Omega_\varepsilon$, we have

$$\begin{aligned} \|\nabla u_\varepsilon\|_{L^2(\Omega_\varepsilon)}^2 &\leq C \int_{\Omega_\varepsilon} \int_{\Omega_\varepsilon} [(\nabla u_\varepsilon(y) - \nabla u_\varepsilon(x))^2 \mu(|y-x|)] \, dy \, dx \quad (18) \\ &+ C \int_{\Omega_\varepsilon} \int_{\Omega^\varepsilon \setminus \Omega_\varepsilon} [(\nabla u_\varepsilon(y) - \nabla u_\varepsilon(x))^2 \mu(|y-x|)] \, dy \, dx + \|\nabla u_\varepsilon\|_{L^2(\Omega^\varepsilon \setminus \Omega_\varepsilon)}^2. \end{aligned}$$

By using Hölder's inequality in (17) and combining it with (18), we obtain

$$\begin{aligned} \|\nabla u_\varepsilon\|_{L^2(\Omega_\varepsilon)}^2 &\leq C \left\{ \|\nabla f_\varepsilon\|_{L^2(\Omega_\varepsilon)}^2 + \|\nabla u_\varepsilon\|_{L^2(\Omega^\varepsilon \setminus \Omega_\varepsilon)}^2 \right. \quad (19) \\ &\left. + \int_{\Omega_\varepsilon} \int_{\Omega^\varepsilon \setminus \Omega_\varepsilon} [(\nabla u_\varepsilon(y) - \nabla u_\varepsilon(x))\mu(|y-x|)] \cdot \nabla u_\varepsilon(x) \, dy \, dx \right\}. \end{aligned}$$

Since $\nabla u_\varepsilon(x) = (u * \nabla \phi_\varepsilon)(x)$ with $\nabla \phi_\varepsilon(x) = \frac{1}{\varepsilon^{n+1}} \nabla \phi\left(\frac{x}{\varepsilon}\right)$, we have by Young's inequality for convolutions that

$$\|\nabla u_\varepsilon\|_{L^2(\Omega^\varepsilon \setminus \Omega_\varepsilon)}^2 \leq \|u\|_{L^1(\Omega^\varepsilon \setminus \Omega_\varepsilon)}^2 \|\nabla \phi_\varepsilon\|_{L^2(\Omega^\varepsilon \setminus \Omega_\varepsilon)}^2.$$

For $f \in L^\infty(\Omega)$, we have $u \in L^\infty(\Omega)$ by Theorem 3, and since $|\Omega^\varepsilon \setminus \Omega_\varepsilon| < C\varepsilon$ (since the boundary of Ω is smooth), the above estimate becomes

$$\|\nabla u_\varepsilon\|_{L^2(\Omega^\varepsilon \setminus \Omega_\varepsilon)}^2 \leq C\varepsilon^2 \|u\|_{L^\infty(\Omega)}^2 \frac{\|\nabla \phi\|_{L^1(\mathbb{R}^n)}^2}{\varepsilon^2} < C.$$

Finally, we estimate the last term on the RHS of (19) by using Young’s inequality for convolutions.

$$\begin{aligned}
 & \left| \int_{\Omega_\varepsilon} \int_{\Omega^\varepsilon \setminus \Omega_\varepsilon} [(\nabla u_\varepsilon(y) - \nabla u_\varepsilon(x))\mu(|y-x|)] \cdot \nabla u_\varepsilon(x) \, dy \, dx \right| \\
 & \leq \left| \int_{\Omega_\varepsilon} \int_{\Omega^\varepsilon \setminus \Omega_\varepsilon} \nabla u_\varepsilon(y) \cdot \nabla u_\varepsilon(x)\mu(|y-x|) \, dy \, dx \right| + \int_{\Omega_\varepsilon} \int_{\Omega^\varepsilon \setminus \Omega_\varepsilon} |\nabla u_\varepsilon(x)|^2 \mu(|y-x|) \, dy \, dx \\
 & \leq \int_{\Omega_\varepsilon} |[(\chi_{\Omega^\varepsilon \setminus \Omega_\varepsilon}(y)\nabla u_\varepsilon) * \mu](x)| \, dx + \int_{\Omega_\varepsilon} \int_{\Omega^\varepsilon \setminus \Omega_\varepsilon} |\nabla u_\varepsilon(x)|^2 \mu(|y-x|) \, dy \, dx \\
 & \leq \eta \|\nabla u_\varepsilon\|_{L^2(\Omega^\varepsilon)} + C(\eta) \|\mu\|_{L^1((\Omega_\varepsilon))} \|\nabla u_\varepsilon\|_{L^2(\Omega^\varepsilon \setminus \Omega_\varepsilon)}^2 \\
 & \qquad \qquad \qquad + \int_{\Omega_\varepsilon} \int_{\Omega^\varepsilon \setminus \Omega_\varepsilon} |\nabla u_\varepsilon(x)|^2 \mu(|y-x|) \, dy \, dx
 \end{aligned}$$

for $\eta > 0$, and where χ_A denotes the characteristic function for the set A . Since $\mu \in L^1(\mathbb{R}^n)$ and $|\Omega^\varepsilon \setminus \Omega_\varepsilon| < C\varepsilon$, we have that

$$J := \int_{\Omega_\varepsilon} \int_{\Omega^\varepsilon \setminus \Omega_\varepsilon} |\nabla u_\varepsilon(x)|^2 \mu(|y-x|) \, dy \, dx < C\varepsilon \|\nabla u_\varepsilon\|_{L^2(\Omega_\varepsilon)}^2$$

uniformly for every $x \in \mathbb{R}^n$. Thus, we can choose ε sufficiently small to absorb the term J defined above in the LHS of (19). Similarly, by choosing η sufficiently small, we obtain

$$\|\nabla u_\varepsilon\|_{L^2(\Omega_\varepsilon)}^2 \leq C, \tag{20}$$

where C is independent of ε . The uniform estimate from (20) shows that given an open set $\Omega' \subset\subset \Omega$, we can extract a subsequence ∇u_{ε_k} that converges weakly to some $h \in L^2(\Omega')$. (Here $\Omega' \subset\subset \Omega$ indicates the closure of Ω' is a compact subset of Ω .) We finally argue that h is the distributional gradient of u . Let $\psi \in C_c(\Omega')$, we have

$$(h, \psi)_{L^2(\Omega')} = \lim_{k \rightarrow \infty} (\nabla u_{\varepsilon_k}, \psi)_{L^2(\Omega')} = - \lim_{k \rightarrow \infty} (u_{\varepsilon_k}, \nabla \psi)_{L^2(\Omega')} = -(u, \nabla \psi)_{L^2(\Omega')},$$

which verifies that $\nabla u = h$. Hence $u \in W^{1,2}(\Omega)$ since C is independent of ε . Since we already proved that $u \in L^\infty(\Omega)$ the conclusion of the theorem follows. \square

Finally, we conclude this section with a theorem that establishes *infinite* regularity of solutions for semilinear problems.

Theorem 5 (Regularity of solutions to semilinear problems). *Suppose that $\mu \in L^1(\mathbb{R}^n)$ satisfies (A1) and that $\|\mu\|_{L^1(\mathbb{R}^n)} = 1$. In addition, assume that $\mu \in W^{s,2}(\mathbb{R}^n)$. Moreover, assume that the mapping*

$$u \mapsto g_x(u) = u + f(x, u)$$

is invertible on \mathbb{R} with the inverse $u \mapsto g_x^{-1}(u) \in C^\infty(\mathbb{R})$ for every $x \in \Omega$. Under these assumptions, if a solution $u \in L^p(\Omega)$, $p > 1$ of (11) exists, then it must satisfy

$$u \in C^\infty(\Omega).$$

Proof. Since $\mu \in W^{s,2}(\mathbb{R}^n)$ and $u \in L^2(\Omega \cup \mathcal{B})$, we have that $u * \mu \in W^{s,2}(\Omega)$. From the equality $u(x) = g_x^{-1}(u * \mu)(x)$ and the smoothness of g_x^{-1} , we have that $u \in W^{s,2}(\Omega)$. We iterate this argument, and using the increased smoothness on u , we will obtain that $u \in W^{2s,2}(\Omega)$, and so on, to finally obtain $u \in W^{\infty,2}(\Omega)$; hence $u \in C^\infty(\Omega)$. □

Remark 3. The strength of the convolution argument that we used above is explored in full generality in the forthcoming paper (Foss et al. 2017).

Regularity for Higher-Order Nonlocal Problems

In Radu et al. (2017) we introduced the nonlocal biharmonic operator

$$\mathcal{B}_\mu[u] := \mathcal{L}_\mu[\mathcal{L}_\mu u]$$

and nonlocal versions of hinged and clamped boundary conditions and showed well-posedness of the coupled nonlocal systems. Moreover, to strengthen the connection between the local and nonlocal theories, we proved that the nonlocal solutions of the hinged and clamped plate systems converge to their respective classical counterparts.

We will focus here on the nonlocal hinged boundary value problem

$$\begin{cases} \mathcal{B}_\mu[u] = f, & x \in \Omega_{2\delta} \\ u = 0, & x \in \Omega \setminus \Omega_\delta \\ \mathcal{L}_\mu[u] = 0, & x \in \Omega_\delta \setminus \Omega_{2\delta} \end{cases} \tag{21}$$

which we showed is well-posed. Note that the parameter $\delta > 0$ determines the thickness of the regions in Ω in which the hinged boundary conditions are imposed; in particular, it is not related to the horizon of μ . We have the following convergence result:

Theorem 6. Let $\Omega \subset \mathbb{R}^n$, $n \geq 2$, be a bounded domain either of class C^2 or convex of class C^1 , and let μ be a nonnegative and integrable kernel that satisfies assumptions (A2) and (A3). Suppose the sequence of positive scalars $\{\delta_n\}_{n=1}^\infty \subset (0, \infty)$ converges to 0 as $n \rightarrow \infty$. For $f \in L^2(\Omega)$, the solutions u_{δ_n} of the nonlocal hinged problems (21) converge in $L^2(\Omega)$ to the weak (variational) solution $u \in W^{2,2}(\Omega) \cap W_0^{1,2}(\Omega)$ of

$$\begin{cases} \Delta^2 u = f, & x \in \Omega \\ u = \Delta u = 0, & x \in \partial\Omega \end{cases} \quad (22)$$

as $n \rightarrow \infty$. Furthermore, if Ω is smooth, e.g., of class C^4 , then u is also in $W^{4,2}(\Omega)$.

Remark 4. The assumptions of μ are more general in Radu et al. (2017); here, we have imposed additional restrictions so we are in position to use our earlier regularity results.

By iterating the regularity result given by Theorem 4, we obtain the following regularity result.

Theorem 7 (Regularity of solutions to the linear nonlocal hinged system). Let $f \in W^{1,2}(\Omega)$, μ a nonnegative and integrable kernel that satisfies assumptions (A2) and (A3). Then there exists a unique solution u to the system (21) such that $u \in W^{1,2}(\Omega) \cap L^\infty(\Omega)$.

Proof. We decompose the system (22) into two second-order nonlocal boundary value problems:

$$\begin{cases} \mathcal{L}_\delta[u] = v, & x \in \Omega_\delta \\ u = 0, & x \in \Omega \setminus \Omega_\delta \end{cases} \quad (23)$$

$$\begin{cases} \mathcal{L}_\delta[v] = f, & x \in \Omega_{2\delta} \\ v = 0, & x \in \Omega_\delta \setminus \Omega_{2\delta} \end{cases} \quad (24)$$

and use Theorem 4 first for (24) to obtain that $v \in W^{1,2}(\Omega) \cap L^\infty(\Omega)$; then with this regularity for v , we apply the theorem to (23) to show that $u \in W^{1,2}(\Omega) \cap L^\infty(\Omega)$. \square

References

- B. Aksoylu, M.L. Parks, Variational theory and domain decomposition for nonlocal problems. *Appl. Math. Comput.* **217**(14), 6498–6515 (2011)
- F. Andreu-Vaillo, J.M. Mazón, J.D. Rossi, J.J. Toledo-Melero, *Nonlocal Diffusion Problems*. Volume 165 of *Mathematical Surveys and Monographs* (American Mathematical Society, Providence/Real Sociedad Matemática Española, Madrid, 2010)

- P.W. Bates, J. Han, The Dirichlet boundary problem for a nonlocal Cahn-Hilliard equation. *J. Math. Anal. Appl.* **311**(1), 289–312 (2005a)
- P.W. Bates, J. Han, The Neumann boundary problem for a nonlocal Cahn-Hilliard equation. *J. Differ. Equ.* **212**(2), 235–277 (2005b)
- E. Berchio, A. Ferrero, F. Gazzola, Structural instability of nonlinear plates modelling suspension bridges: mathematical answers to some long-standing questions. *Nonlinear Anal. Real World Appl.* **28**, 91–125 (2016)
- L.A. Caffarelli, R. Leitão, J.M. Urbano, Regularity for anisotropic fully nonlinear integro-differential equations. *Math. Ann.* **360**(3–4), 681–714 (2014)
- Q. Du, M. Gunzburger, R.B. Lehoucq, K. Zhou, Analysis and approximation of nonlocal diffusion problems with volume constraints. *SIAM Rev.* **54**(4), 667–696 (2012)
- Q. Du, M. Gunzburger, R.B. Lehoucq, K. Zhou, A nonlocal vector calculus, nonlocal volume-constrained problems, and nonlocal balance laws. *Math. Models Methods Appl. Sci.* **23**(03), 493–540 (2013)
- M. Foss, J. Geisbauer, Partial regularity for subquadratic parabolic systems with continuous coefficients. *Manuscripta Math.* **139**(1–2), 1–47 (2012)
- M. Foss, P. Radu, Differentiability and integrability properties for solutions to nonlocal equations, in *New Trends in Differential Equations, Control Theory and Optimization: Proceedings of the 8th Congress of Romanian Mathematicians* (World Scientific, 2016), pp. 105–119
- M. Foss, P. Radu, C. Wright, Regularity and existence of minimizers for nonlocal energy functionals. *Differ. Integr. Equ.* (2017, to appear)
- F. Gazzola, *Mathematical Models for Suspension Bridges: Nonlinear Structural Instability*. Volume 15 of MS&A. Modeling, Simulation and Applications (Springer, Cham, 2015)
- G. Gilboa, S. Osher, Nonlocal operators with applications to image processing. *Multiscale Model. Simul.* **7**(3), 1005–1028 (2008)
- B. Hinds, P. Radu, Dirichlet’s principle and wellposedness of solutions for a nonlocal p -Laplacian system. *Appl. Math. Comput.* **219**(4), 1411–1419 (2012)
- S. Mayboroda, V. Maz’ya, Regularity of solutions to the polyharmonic equation in general domains. *Invent. Math.* **196**(1), 1–68 (2014)
- T. Mengesha, Q. Du, The bond-based peridynamic system with Dirichlet-type volume constraint. *Proc. R. Soc. Edinb. Sect. A* **144**(1), 161–186 (2014)
- A. Mogilner, L. Edelstein-Keshet, A non-local model for a swarm. *J. Math. Biol.* **38**(6), 534–570 (1999)
- S. Oterkus, E. Madenci, A. Agwai, Peridynamic thermal diffusion. *J. Comput. Phys.* **265**, 71–96 (2014)
- P. Radu, D. Toundykov, J. Trageser, Finite time blow-up in nonlinear suspension bridge models. *J. Differ. Equ.* **257**(11), 4030–4063 (2014)
- P. Radu, D. Toundykov, J. Trageser, A nonlocal biharmonic operator and its connection with the classical analogue. *Arch. Ration. Mech. Anal.* **223**(2), 845–880 (2017)
- P. Radu, K. Wells, A state-based Laplacian: properties and convergence to its local and nonlocal counterparts (2017, Preprint)
- S. Silling, Reformulation of elasticity theory for discontinuities and long-range forces. *J. Mech. Phys. Solids* **48**, 175–209 (2000)
- H. Sun, D. Uminsky, A.L. Bertozzi, Stability and clustering of self-similar solutions of aggregation equations. *J. Math. Phys.* **53**(11), 115610, 18 (2012)
- X. Tian, Q. Du, Asymptotically compatible schemes and applications to robust discretization of nonlocal models. *SIAM J. Numer. Anal.* **52**(4), 1641–1665 (2014)



Dynamic Brittle Fracture from Nonlocal Double-Well Potentials: A State-Based Model

38

Robert Lipton, Eyad Said, and Prashant K. Jha

Contents

Introduction	1266
Nonlocal Dynamics	1267
Existence of Solutions	1270
Stability Analysis	1273
Control of the Softening Zone	1276
Calibration of the Model	1281
Calibrating the Peridynamic Energy to Elastic Properties	1281
Calibrating Energy Release Rate	1286
Linear Elastic Operators in the Limit of Vanishing Horizon	1288
Conclusions	1290
References	1291

Abstract

We introduce a regularized model for free fracture propagation based on nonlocal potentials. We work within the small deformation setting, and the model is developed within a state-based peridynamic formulation. At each instant of the evolution, we identify the softening zone where strains lie above the strength of the material. We show that deformation discontinuities associated with flaws larger than the length scale of nonlocality δ can become unstable and grow. An explicit inequality is found that shows that the volume of the softening zone

R. Lipton (✉)

Department of Mathematics and Center for Computation and Technology, Louisiana State University, Baton Rouge, LA, USA
e-mail: lipton@lsu.edu

E. Said · P. K. Jha

Department of Mathematics, Louisiana State University, Baton Rouge, LA, USA
e-mail: esaid1@lsu.edu; prashant.j16o@gmail.com

goes to zero linearly with the length scale of nonlocal interaction. This scaling is consistent with the notion that a softening zone of width proportional to δ converges to a sharp fracture set as the length scale of nonlocal interaction goes to zero. Here the softening zone is interpreted as a regularization of the crack network. Inside quiescent regions with no cracks or softening, the nonlocal operator converges to the local elastic operator at a rate proportional to the radius of nonlocal interaction. This model is designed to be calibrated to measured values of critical energy release rate, shear modulus, and bulk modulus of material samples. For this model one is not restricted to Poisson ratios of 1/4 and can choose the potentials so that small strain behavior is specified by the isotropic elasticity tensor for any material with prescribed shear and Lamé moduli.

Keywords

Free fracture model · Nonlocal interactions · Double-well potentials · State-based peridynamics

Introduction

We address the problem of free crack propagation in homogeneous materials. The crack path is not known a priori and is found as part of the problem solution. Our approach is to use a nonlocal formulation based on double-well potentials. We will work within the small deformation setting, and the model is developed within a state-based peridynamic formulation. Peridynamics (Silling 2000; Silling et al. 2007) is a nonlocal formulation of continuum mechanics expressed in terms of displacement differences as opposed to spatial derivatives of the displacement field. These features provide the ability to simultaneously simulate both smooth displacements and defect evolution. Computational methods based on peridynamic modeling exhibit formation and evolution of sharp features associated with phase transformation (see Dayal and Bhattacharya 2006) and fracture (see Silling and Lehoucq 2008; Silling et al. 2010; Foster et al. 2011; Agwai et al. 2011; Du et al. 2013; Lipton et al. 2016; Bobaru and Hu 2012; Ha and Bobaru 2010; Silling and Bobaru 2005; Weckner and Abeyaratne 2005; Gerstle et al. 2007; Silling and Askari 2005). A recent review of the state of the art can be found in Bobaru et al. (2016).

In this work we are motivated by the recent models proposed and studied in Lipton (2014, 2016), and Lipton et al. (2016). Calibration has been investigated in Diehl et al. (2016). These models are defined by double-well two-point strain potentials. Here one potential well is centered at the origin and associated with elastic response, while the other well is at infinity and associated with surface energy. The rationale for studying these models is that they are shown to be well posed over the class of square-integrable non-smooth displacements, and in the limit of vanishing nonlocality, the dynamics localize and recover features of sharp fracture propagation (see Lipton 2014, 2016). In this work we extend this modeling approach to the state-based formulation. Our work is further motivated by the recent numerical-experimental study carried out in Diehl et al. (2016) demonstrating that

the bond-based model is unable to capture the Poisson ratio for a sample of PMMA at room temperature. Here we develop a double-well state-based potential for which the Poisson ratio is no longer constrained to be $1/4$. We show that for this model we can choose the potentials so that the small strain behavior is specified by the isotropic elasticity tensor for any material with prescribed shear and Lamé moduli.

Nonlocal Dynamics

We formulate the nonlocal dynamics. Here we will assume displacements u are small (infinitesimal) relative to the size of the three-dimensional body D . The tensile strain is written as $S = S(y, x, t; u)$ and given by

$$S(y, x, t; u) = \frac{u(t, y) - u(t, x)}{|y - x|} \cdot e_{y-x}, \quad e_{y-x} = \frac{y - x}{|y - x|}, \quad (1)$$

where e_{y-x} is a unit direction vector and \cdot is the dot product. It is evident that $S(y, x, t; u)$ is the tensile strain along the direction e_{y-x} . We introduce the influence function $\omega^\delta(|y - x|)$ such that ω^δ is nonzero for $|y - x| < \delta$, zero outside. Here we will take $\omega^\delta(|y - x|) = \omega(|y - x|/\delta)$ with $\omega(r) = 0$ for $r > 1$ nonnegative for $r < 1$ and ω is bounded.

The spherical or hydrostatic strain at x is given by

$$\theta(x, t; u) = \frac{1}{V_\delta} \int_{D \cap B_\delta(x)} \omega^\delta(|y - x|) S(y, x, t; u) |y - x| dy, \quad (2)$$

where V_δ is the volume of the ball $B_\delta(x)$ of radius δ centered at x . Here we have employed the normalization $|y - x|/\delta$ so that this factor takes values in the interval from 0 to 1.

Motivated by potentials of Lennard-Jones type, we define the force potential for tensile strain given by

$$\mathcal{W}^\delta(S(y, x, t; u)) = \alpha \omega^\delta(|y - x|) \frac{1}{\delta |y - x|} f(\sqrt{|y - x|} S(y, x, t; u)) \quad (3)$$

and the potential for hydrostatic strain

$$\mathcal{V}^\delta(\theta(x, t; u)) = \frac{\beta g(\theta(x, t; u))}{\delta^2} \quad (4)$$

where $\mathcal{W}^\delta(S(y, x, t; u))$ is the pairwise force potential per unit length between two points x and y and $\mathcal{V}^\delta(\theta(x, t; u))$ is the hydrostatic force potential density at x . They are described in terms of their potential functions f and g (see Fig. 1). These two potentials are double-well potentials that are chosen so that the associated forces acting between material points x and y are initially elastic and then soften and

Fig. 1 Potential function $f(r)$ for tensile force and potential function $g(r)$ for hydrostatic force

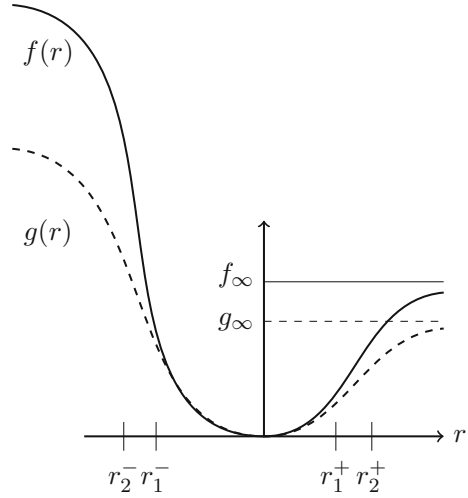
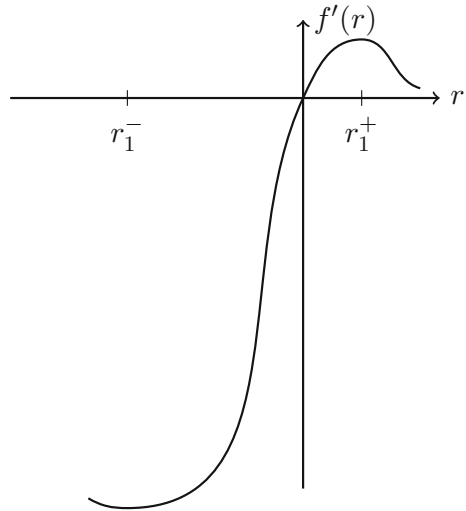


Fig. 2 Cohesive tensile force



decay to zero as the strain between points increases (see Fig. 2 for the tensile force). This force is negative for compression, and a similar force hydrostatic strain law follows from the potential for hydrostatic strain. The first well for $\mathcal{W}^\delta(S(y, x, t; u))$ and $\mathcal{V}^\delta(\theta(x, t; u))$ is at zero tensile and hydrostatic strain, respectively. With this in mind, we make the choice

$$f(0) = f'(0) = g(0) = g'(0). \tag{5}$$

The second well is at infinite tensile and hydrostatic strain and is characterized by the horizontal asymptotes $\lim_{S \rightarrow \infty} f(S) = f_\infty$ and $\lim_{\theta \rightarrow \infty} g(\theta) = g_\infty$, respectively (see Fig. 1).

The critical tensile strain $S_c^+ > 0$ for which the force begins to soften is given by the inflection point $r_1^+ > 0$ of f and is

$$S_c^+ = \frac{r_1^+}{\sqrt{|y-x|}}. \tag{6}$$

The critical negative tensile strain is chosen much larger in magnitude than S_c^+ and is

$$S_c^- = \frac{r_1^-}{\sqrt{|y-x|}}, \tag{7}$$

with $r_1^- < 0$ and $r_1^+ \ll |r_1^-|$. The critical value $0 < \theta_c^+$ where the force begins to soften under positive hydrostatic strain for $\theta(x, t; u) > \theta_c^+$ is given by the inflection point r_2^+ of g and is

$$\theta_c^+ = r_2^+. \tag{8}$$

The critical compressive hydrostatic strain where the force begins to soften for negative hydrostatic strain is chosen much larger in magnitude than θ_c^+ and is

$$\theta_c^- = r_2^-, \tag{9}$$

with $r_2^- < 0$ and $r_2^+ < |r_2^-|$. For this model we suppose the inflection points for g and f satisfy the ordering

$$r_2^- < r_1^- < 0 < r_1^+ < r_2^+. \tag{10}$$

This ordering is chosen to illustrate ideas for a material that is weaker in shear strain than hydrostatic strain. With this choice and the appropriate influence function ω^δ , if the hydrostatic stress is positive at x and is above the critical value θ_c^+ , then there are points y in the peridynamic neighborhood for which the tensile stress between x and y is above S_c^+ . This aspect of the model is established and addressed in section “Control of the Softening Zone.”

The potential energy is given by

$$PD^\delta(u) = \frac{1}{V_\delta} \int_D \int_{D \cap B_\delta(x)} |y-x| \mathcal{W}^\delta(S(y, x, t; u)) dy dx + \int_D \mathcal{V}^\delta(\theta(x, t; u)) dx. \tag{11}$$

The material is assumed homogeneous, and the density is given by ρ , and the applied body force is denoted by $b(x, t)$. We define the Lagrangian

$$L(u, \partial_t u, t) = \frac{\rho}{2} \|\dot{u}\|_{L^2(D; \mathbb{R}^3)}^2 - PD^\delta(u) + \int_D b \cdot u dx,$$

here \dot{u} is the velocity given by the time derivative of u , and $\|\dot{u}\|_{L^2(D; \mathbb{R}^3)}$ denotes the L^2 norm of the vector field $\dot{u} : D \rightarrow \mathbb{R}^3$. Applying the principle of least action together with a straightforward calculation gives the nonlocal dynamics

$$\rho \ddot{u}(x, t) = \mathcal{L}^T(u)(x, t) + \mathcal{L}^D(u)(x, t) + b(x, t), \text{ for } x \in D, \quad (12)$$

where

$$\mathcal{L}^T(u)(x, t) = \frac{2\alpha}{V_\delta} \int_{D \cap B_\delta(x)} \frac{\omega^\delta(|y-x|)}{\delta|y-x|} \partial_S f(\sqrt{|y-x|} S(y, x, t; u)) e_{y-x} dy, \quad (13)$$

and

$$\mathcal{L}^D(u)(x, t) = \frac{\beta}{V_\delta} \int_{D \cap B_\delta(x)} \frac{\omega^\delta(|y-x|)}{\delta^2} [\partial_\theta g(\theta(y, t; u)) + \partial_\theta g(\theta(x, t; u))] e_{y-x} dy. \quad (14)$$

The dynamics is complemented with the initial data

$$u(x, 0) = u_0(x), \quad \partial_t u(x, 0) = v_0(x). \quad (15)$$

It is readily verified that this is an ordinary state-based peridynamic model. The forces are defined by the derivatives of the potential functions, and the derivative associated with the tensile strain potential is sketched in Fig. 2. We show in the next section that this initial value problem is well posed.

Existence of Solutions

The regularity and existence of the solution depends on the regularity of the initial data and body force. In this work we choose a general class of body forces and initial conditions. The initial displacement u_0 and velocity v_0 are chosen to be integrable and belonging to $L^\infty(D; \mathbb{R}^3)$. The body force $b(x, t)$ is chosen such that for every $t \in [0, T_0]$, b takes values in $L^\infty(D, \mathbb{R}^3)$ and is continuous in time. The associated norm is defined to be $\|b\|_{C([0, T_0]; L^\infty(D, \mathbb{R}^3))} = \max_{t \in [0, T_0]} \|b(x, t)\|_{L^\infty(D, \mathbb{R}^3)}$. The space of continuous functions in time taking values in $L^\infty(D; \mathbb{R}^3)$ for which this norm is finite is denoted by $C([0, T_0]; L^\infty(D, \mathbb{R}^3))$. The space of functions twice

differentiable in time taking values in $L^\infty(D, \mathbb{R}^3)$ such that both derivatives belong to $C([0, T_0]; L^\infty(D, \mathbb{R}^3))$ is written as $C^2([0, T_0]; L^\infty(D, \mathbb{R}^3))$.

We will establish existence and uniqueness for the evolution by writing the second-order ODE as an equivalent first-order system. The nonlocal dynamics (12) can be written as a first-order system. Set $y = (y_1, y_2)$ where $y_1 = u$ and $y_2 = u_t$. Now, set $F^\delta(y, t) = (F_1(y, t), F_2(y, t))^T$ where:

$$\begin{aligned} F_1(y, t) &= y_2 \\ F_2(y, t) &= \mathcal{L}^T(y_1)(t) + \mathcal{L}^D(y_1)(t) + b(t) \end{aligned} \tag{16}$$

And the initial value problem is given by the equivalent first-order system

$$\begin{aligned} \frac{d}{dt}y^\delta &= F^\delta(y^\delta, t) \\ y(0) &= (y_1(0), y_2(0)) = (u_0, v_0) \end{aligned} \tag{17}$$

The existence of a unique solution to the initial value problem is asserted in the following theorem.

Theorem 1. *For a body force $b(t, x)$ in $C^1([0, T]; L^\infty(D, \mathbb{R}^3))$ and initial data $y_1(0)$ and $y_2(0)$ in $L^\infty_0(D; \mathbb{R}^3) \times L^\infty_0(D; \mathbb{R}^3)$, there exists a unique solution $y(t)$ such that $y_1 = u$ is in $C^2([0, T]; L^\infty(D, \mathbb{R}^3))$ for the dynamics described by (17) with initial data in $L^\infty(D; \mathbb{R}^3) \times L^\infty(D; \mathbb{R}^3)$ and body force $b(t, x)$ in $C^1([0, T]; L^\infty(D, \mathbb{R}^3))$.*

Proof. We will show that the model is Lipschitz continuous and then apply the theory of ODE in Banach spaces, e.g., Driver (2003), to guarantee the existence of a unique solution. It is sufficient to show that

$$\|\mathcal{L}^T(u)(x, t) + \mathcal{L}^D(u)(x, t) - (\mathcal{L}^T(v)(x, t) + \mathcal{L}^D(v)(x, t))\|_{L^\infty(D)} \leq C\|u - v\|_{L^\infty(D)} \tag{18}$$

For ease of notation, we introduce the following vectors

$$\begin{aligned} \vec{U} &= u(y) - u(x), \\ \vec{V} &= v(y) - v(x). \end{aligned}$$

We write

$$\mathcal{L}^T(u)(x, t) + \mathcal{L}^D(u)(x, t) - (\mathcal{L}^T(v)(x, t) + \mathcal{L}^D(v)(x, t)) = \mathcal{I}_1 + \mathcal{I}_2. \tag{19}$$

Here

$$\begin{aligned}
 \mathcal{I}_1 &= \frac{2\alpha}{\delta V_\delta} \int_{D \cap B_\delta(x)} \frac{\omega^\delta(|y-x|)}{\sqrt{|y-x|}} \left\{ f'(\sqrt{|y-x|}S(y,x,t;u)) \right. \\
 &\quad \left. - f'(\sqrt{|y-x|}S(y,x,t;v)) \right\} e_{(y-x)} dy \\
 \mathcal{I}_2 &= \frac{\beta}{\delta^2 V_\delta} \int_{D \cap B_\delta(x)} \omega^\delta(|y-x|) \left(g'(\theta(y,t;u)) + g'(\theta(x,t;u)) \right. \\
 &\quad \left. - (g'(\theta(y,t;v)) + g'(\theta(x,t;v))) \right) e_{(y-x)} dy
 \end{aligned} \tag{20}$$

Since f'' is bounded a straightforward calculation gives:

$$\begin{aligned}
 &|f'(\sqrt{|y-x|}S(y,x,t;u)) - f'(\sqrt{|y-x|}S(y,x,t;v))| \\
 &\leq \sqrt{|y-x|} \sup_{s \in \mathbb{R}} \{|f''(s)|\} |S(y,x,t;u) - S(y,x,t;v)|,
 \end{aligned}$$

and $|e_{y-x}| = 1$, so we can bound \mathcal{I}_1 by

$$|\mathcal{I}_1| \leq \frac{2\alpha}{\delta V_\delta} \int_{D \cap B_\delta(x)} \omega^\delta(|y-x|) \sup_{x \in D} \{|f''(x)|\} |S(y,x,t;u) - S(y,x,t;v)| dy. \tag{21}$$

In what follows $C_1 = \sup_{s \in \mathbb{R}} \{|f''(s)|\} < \infty$ and we make the change of variable

$$\begin{aligned}
 y &= x + \delta \xi \\
 |y-x| &= \sigma |\xi| \\
 dy &= \delta^3 d\xi,
 \end{aligned}$$

and a straightforward calculation shows

$$\mathcal{I}_1 \leq \frac{2\alpha C_1}{\delta^2} \int_{H_1(0) \cap \{x+\delta\xi \in D\}} |\omega(\xi)| \frac{|u(x+\delta\xi) - u(x) - (v(x+\delta\xi) - v(x))|}{|\xi|} d\xi \tag{22}$$

Which leads to the inequality

$$\|\mathcal{I}_1\|_{L^\infty(D; \mathbb{R}^3)} \leq \frac{4\alpha C_1 C_2}{\delta^2} \|u - v\|_{L^\infty(D; \mathbb{R}^3)}, \tag{23}$$

with $C_2 = \int_{H_1(0)} |\xi|^{-1} \omega(|\xi|) d\xi$. Now we can work on the second part, where we follow a similar approach. Noting that g'' is bounded, we let $C_3 = \sup_{\theta \in \mathbb{R}} \{|g''(\theta)|\} < \infty$ and $C_4 = \int_{H_1(0)} |\xi| \omega(|\xi|) d\xi$, to find that

$$\begin{aligned}
 |g'(\theta(y, t; u)) - g'(\theta(y, t; v))| &\leq C_3|\theta(y, t; u) - \theta(y, t; v)| \\
 &\leq \frac{2C_3C_4}{\delta^2} \|u - v\|_{L^\infty(D; \mathbb{R}^3)},
 \end{aligned}$$

and

$$\begin{aligned}
 |g'(\theta(x, t; u)) - g'(\theta(x, t; v))| &\leq C_3|\theta(x, t; u) - \theta(x, t; v)| \\
 &\leq \frac{2C_3C_4}{\delta^2} \|u - v\|_{L^\infty(D; \mathbb{R}^3)},
 \end{aligned}$$

so

$$\|\mathcal{I}_2\|_{L^\infty(D; \mathbb{R}^3)} \leq \frac{4\beta C_3C_4}{\delta^2} \|u - v\|_{L^\infty(D; \mathbb{R}^3)}. \tag{24}$$

Adding (23) and (24) gives the desired result

$$\begin{aligned}
 &\| \mathcal{L}^T(u)(x, t) + \mathcal{L}^D(u)(x, t) - (\mathcal{L}^T(v)(x, t) + \mathcal{L}^D(v)(x, t)) \|_{L^\infty(D; \mathbb{R}^3)} \\
 &\leq \frac{4(\alpha C_1C_2 + \beta C_3C_4)}{\delta^2} \|u - v\|_{L^\infty(D; \mathbb{R}^3)}.
 \end{aligned} \tag{25}$$

□

Stability Analysis

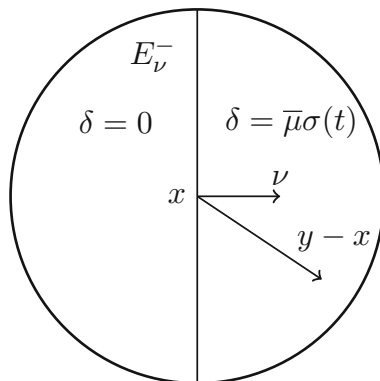
In this section we identify a source for crack nucleation as a material defect represented by a jump discontinuity in the displacement field. To illustrate the ideas, we assume the defect is in the interior of the body and at least δ away from the boundary. This jump discontinuity can become unstable and grow in time. We proceed with a perturbation analysis and consider a time-independent body force density b and a smooth equilibrium solution u . Now assume that the defect perturbs u in the neighborhood of a point x by a piecewise constant vector field s that represents a jump in displacement across a planar surface with normal vector v . We assume that this jump occurs along a defect of length 2δ on the planar surface.

The smooth equilibrium solution $u(x, t)$ is a solution of

$$0 = \mathcal{L}^T(u)(x, t) + \mathcal{L}^D(u)(x, t) + b(x) \tag{26}$$

Now consider a perturbed solution $u^P(x, t)$ that differs from equilibrium solution $u(x, t)$ by the jump across the planar surface which is specified by unit normal vector v . We suppose the surface passes through x and extends across the peridynamic neighborhood centered at x . Points y for which $(y - x) \cdot v < 0$ are denoted by \mathcal{E}_v^- and points for which $(y - x) \cdot v \geq 0$ are denoted by \mathcal{E}_v^+ , see Fig. 3.

Fig. 3 Jump discontinuity



The perturbed solution u^P satisfies

$$\rho \ddot{u}^P = \mathcal{L}^T(u^P)(x, t) + \mathcal{L}^D(u^P)(x, t) + b(x) \tag{27}$$

Here the perturbed solution $u^P(x, t)$ is given by the equilibrium solution plus a piecewise constant perturbation and is written

$$u^P(y, t) = u(y, t) + s(y, t) \tag{28}$$

Where

$$s(y, t) = \begin{cases} 0 & y \in \mathcal{E}_v^- \\ \bar{\mu}\sigma(t) & y \in \mathcal{E}_v^+ \end{cases} \tag{29}$$

Subtracting (26) from (27) gives

$$\rho \ddot{u}^P = \mathcal{L}^T(u^P)(x, t) + \mathcal{L}^D(u^P)(x, t) - \mathcal{L}^T(u)(x, t) + \mathcal{L}^D(u)(x, t) \tag{30}$$

Here the second term in $\mathcal{L}^D(u)$ vanishes as we are away from the boundary, and the integrand is odd in the y variable with respect to the domain $B_\delta(x)$. Since $u^P = u + s$ and s is small, we expand $f'(\sqrt{|y-x|}(S(y, x, t; u + s)))$ in Taylor series in s . Noting that $\theta(x, t; u + s) = \theta(x, t; u) + \theta(x, t; s)$ and $\theta(x, t; s)$ is initially infinitesimal, we also expand $g'(\theta(x, t; u + s))$ in a Taylor series in $\theta(x, t; s)$. Applying the expansions to (30) shows that to leading order

$$\begin{aligned} \rho \ddot{u}^P &= \rho \ddot{\sigma} \bar{\mu} = \frac{2\alpha}{\delta V_\delta} \int_{B_\delta(x)} \frac{\omega^\delta(|y-x|)}{\sqrt{|y-x|}} f''(\sqrt{|y-x|}S)(s(y, t) - s(x, t)) \cdot \\ &e_{(y-x)} e_{(y-x)} dy + \frac{\beta}{V_\delta \delta^2} \int_{B_\delta(x)} \omega^\delta(|y-x|) g''(\theta(y, t; u)) \frac{1}{V_\delta} \\ &\int_{B_\delta(y)} \omega^\delta(|z-y|) (s(z, t) - s(y, t)) \cdot e_{z-y} dz e_{y-x} dy = I_1 + I_2, \end{aligned} \tag{31}$$

where I_1 and I_2 are the first, second terms on the right-hand side of (31). A straightforward calculation using (29) shows that

$$I_1 = -\frac{2\alpha}{\delta V_\delta} \int_{B_\delta(x) \cap \mathcal{E}_v^-} \frac{J^\delta(|y-x|)}{|y-x|} f''(\sqrt{|y-x|}S) e_{(y-x)} \cdot \bar{\mu}\sigma(t) e_{(y-x)} dy \quad (32)$$

We next calculate I_2 . A straightforward but delicate calculation gives

$$\frac{1}{V_\delta} \int_{B_\delta(y)} \frac{\omega^\delta(|z-y|)}{\delta} (s(z,t) - s(y,t)) \cdot e_{z-y} dz = b(y) \cdot \bar{\mu}\sigma(t) \quad (33)$$

where

$$b(y) = \frac{1}{V_\delta} \int_0^{2\pi} \int_a^\delta \int_0^{\bar{\phi}} \omega(|z-y|) e(\theta, \phi) |z-y|^2 \sin \phi d\phi d\theta d|z-y| \quad (34)$$

and the limits of the iterated integral are

$$a = |(y-x) \cdot v| \quad \bar{\phi} = \arccos \left(\frac{|(y-x) \cdot v|}{|z-y|} \right), \quad (35)$$

and $e(\theta, \phi)$ is the vector on the unit sphere with direction specified by the angles θ and ϕ . Calculation now gives

$$I_2 = \frac{\beta}{V_\delta \delta^2} \left(\int_{B_\delta(x) \cap \mathcal{E}_v^-} \omega^\delta(|y-x|) g''(\theta(y,t;u)) b(y) \cdot \bar{\mu}\sigma(t) e_{y-x} dy - \int_{B_\delta(x) \cap \mathcal{E}_v^+} \omega^\delta(|y-x|) g''(\theta(y,t;u)) b(y) \cdot \bar{\mu}\sigma(t) e_{y-x} dy \right), \quad (36)$$

where

$$\begin{aligned} & \int_{B_\delta(x) \cap \mathcal{E}_v^-} \frac{\omega^\delta(|y-x|)}{\delta} b(y) \cdot \bar{\mu}s(t) e_{y-x} dy \\ &= \int_{B_\delta(x) \cap \mathcal{E}_v^+} \frac{\omega^\delta(|y-x|)}{\delta} b(y) \cdot \bar{\mu}s(t) e_{y-x} dy. \end{aligned} \quad (37)$$

We now take the dot product of both sides of (31) with $\bar{\mu}$ to get

$$\rho \ddot{\sigma} = \frac{(\mathbf{A} + \mathbf{B}^{\text{sym}}) \bar{\mu} \cdot \bar{\mu}}{|\bar{\mu}|^2} \sigma(t), \quad (38)$$

where

$$\mathbf{A} = -\frac{2\alpha}{\delta V_\delta} \int_{B_\delta(x) \cap \mathcal{E}_v^-} \frac{J^\delta(|y-x|)}{|y-x|} f''(\sqrt{|y-x|S}) e_{(y-x)} \otimes e_{(y-x)} dy \quad (39)$$

and $\mathbf{B}^{\text{sym}} = (\mathbf{B} + \mathbf{B}^T)/2$ with

$$\begin{aligned} \mathbf{B} = & \frac{\beta}{V_\delta \delta^2} \left(\int_{B_\delta(x) \cap \mathcal{E}_v^-} \omega^\delta(|y-x|) g''(\theta(y,t;u)) b(y) \otimes e_{y-x} dy \right. \\ & \left. - \int_{B_\delta(x) \cap \mathcal{E}_v^+} \omega^\delta(|y-x|) g''(\theta(y,t;u)) b(y) \otimes e_{y-x} dy \right). \end{aligned} \quad (40)$$

Inspection shows that

$$f''(\sqrt{|y-x|S}) < 0, \text{ when } S > S_c^+. \quad (41)$$

Thus the eigenvalues of \mathbf{A} can be nonnegative whenever the tensile strain is positive and greater than S_c^+ so that the force is in the softening regime for a preponderance of points y inside $B_\delta(x)$. In general the defect will be stable if all eigenvalues of the stability matrix $\mathbf{A} + \mathbf{B}^{\text{sym}}$ are negative. On the other hand, the defect will be unstable if at least one eigenvalue of the stability matrix is positive.

We collect results in the following proposition.

Proposition 1 (Fracture nucleation condition about a defect). *A condition for crack nucleation at a defect passing through a point x is that the associated stability matrix $\mathbf{A} + \mathbf{B}^{\text{sym}}$ has at least one positive eigenvalue.*

If the equilibrium solution is constant, then $\theta(y,t;u) = \text{constant}$ and $\mathbf{B}^{\text{sym}} = 0$. For this case the fracture nucleation condition simplifies and depends only on the eigenvalues of the matrix \mathbf{A} . In the next section, we analyze the size of the set where the tensile strain is greater than S_c^+ so that the tensile force is in the softening regime for points y inside $B_\delta(x)$.

Control of the Softening Zone

We define the softening zone in terms of the collection of centers of peridynamic neighborhoods with tensile strain exceeding S_c^+ . In what follows we probe the dynamics to obtain mathematically rigorous and explicit estimates on the size of the softening zone in terms of the radius of the peridynamic horizon. In this section we assume $\omega^\delta = 1$, $\delta < 1$, and from the definition of the hydrostatic strain $\theta(x,t;u)$, we have the following lemma.

Lemma 1 (Hydrostatic softening implies tensile softening). *If $\theta_c^+ < \theta(x, t; u)$, then $S_c^+ < S(y, x, t; u)$ for some subset of points y inside the peridynamic neighborhood centered at x and*

$$\{x \in D : \theta(x) > \theta_c^+\} \subset \{x \in D : S(x, y, t; u) > S_c^+, \text{ for some } y \text{ in } B_\delta(x)\}. \quad (42)$$

Proof. Suppose $\theta_c^+ < \theta(x, t; u)$, then there are points y in $B_\delta(x)$ for which

$$\theta_c^+ < |y - x|S(y, x, t; u) < \sqrt{|y - x|}S(y, x, t; u), \quad (43)$$

so

$$S_c^+ < \frac{\theta_c^+}{\sqrt{|y - x|}} < S(y, x, t; u), \quad (44)$$

since $r_1^+ < r_2^+ = \theta_c^+$. This directly implies

$$\{x \in D : \theta(x) > \theta_c^+\} \subset \{x \in D : S(x, y, t; u) > S_c^+, \text{ for some } y \text{ in } B_\delta(x)\}, \quad (45)$$

and the lemma is proved. \square

This inequality shows that the collection of neighborhoods where softening is due to the hydrostatic force is also subset of the neighborhoods where there is softening due to tensile force. Motivated by this observation, we focus on peridynamic neighborhoods where the tensile strain is above critical. We start by defining the *softening zone*. The set of points y in $B_\delta(x)$ with tensile strain larger than critical can be written as

$$A_\delta^+(x) = \{y \in B_\delta(x) : S(y, x, t; u) > S_c^+\}.$$

From the monotonicity of the force potential f , we can also express this set as

$$A_\delta^+(x) = \{y \in B_\delta(x); f(\sqrt{|y - x|}S(y, x, t; u)) \geq f(r_1^+)\}.$$

We define the weighted volume of the set A_δ^+ in terms of its characteristic function $\chi_{A_\delta^+}(y)$ taking the value one for $y \in A_\delta^+$ and zero outside. The weighted volume of A_δ^+ is given by $\int_{B_\delta(x)} \chi_{A_\delta^+}(y)|y - x| dy$, and the weighted volume of $B_\delta(x)$ is $m = \int_{B_\delta(x)} |y - x| dy$. The weighted volume fraction $P_\delta(x)$ of $y \in B_\delta(x)$ with tensile strain larger than critical is given by the ratio

$$P_\delta(x) = \frac{\int_{B_\delta(x)} \chi_{A_\delta^+}(y)|y - x| dy}{m}.$$

Definition 1 (Softening zone). Fix any volume fraction $0 < \gamma \leq 1$, and with each time t in the interval $0 \leq t \leq T$, define the softening zone $SZ^\delta(\gamma, t)$ to be the collection of centers of peridynamic neighborhoods for which the weighted volume fraction of points y with strain $S(y, x, t; u)$ exceeding the threshold S_c is greater than γ , i.e.,

$$SZ^\delta(\gamma, t) = \{x \in D; P_\delta(x) > \gamma\}. \tag{46}$$

We now show that the volume of $SZ^\delta(\gamma, t)$ goes to zero linearly with the horizon δ for properly chosen initial data and body force. This scaling is consistent with the notion that a softening zone of width proportional to δ converges to a sharp fracture as the length scale δ of nonlocal interaction goes to zero. We define the sum of kinetic and potential energy as

$$W(t) = \frac{\rho}{2} \|\dot{u}\|_{L^2(D, \mathbb{R}^d)}^2 + PD^\delta(u(t)) \tag{47}$$

and set

$$C(t) = \left(\frac{1}{\sqrt{\rho}} \int_0^t \|b\|_{L^2(D, \mathbb{R}^d)} d\tau + \sqrt{W(0)} \right)^2. \tag{48}$$

Here $C(t)$ is a measure of the total energy delivered to the body from initial conditions and body force up to time t . The tensile toughness is defined to be the energy of tensile tension between x and y per unit length necessary for softening and is given by $f(r_1^+)/\delta$. We now state the geometric dependence of the softening zone on horizon.

Theorem 2. *The volume of the softening zone SZ^δ is controlled by the horizon δ according to the following relation expressed in terms of the total energy delivered to the system, the tensile toughness, and the weighted volume fraction of points y where the tensile strain exceeds S_c^+ ,*

$$\text{Volume}(SZ^\delta(\gamma, t)) \leq \frac{\delta C(t)}{\gamma m f(r_1^+)}. \tag{49}$$

Remark 1. It is clear that for zero initial data such that $u(0, x) = 0$ that $C(t)$ depends only on the body force $b(t, x)$ and initial velocity. For this choice we see that the softening zone goes to zero linearly with the horizon δ .

We now establish the theorem using Gronwall’s inequality and Tchebychev’s inequality. The peridynamic energy density at a point x is

$$E^\delta(x) = \frac{1}{V_\delta} \int_{D \cap B_\delta(x)} |y - x| \mathcal{W}^\delta(S(y, x, t; u)) dy + \mathcal{V}^\delta(\theta(x, t; u)) \tag{50}$$

Which can also be rewritten with the following change of variable $y - x = \delta\xi$

$$E^\delta(x) = \frac{\alpha}{\delta V_1} \int_{D \cap B_1(0)} \omega(|\xi|) f(\sqrt{\delta|\xi|} S(x + \delta\xi, x, t; u)) d\xi + \frac{\beta g(\theta(x, t; u))}{\delta^2} \tag{51}$$

Recall from the monotonicity of $f(r)$ that $r_1^+ < r$ implies $f(r_1^+) < f(r)$. Now define the set where the strain exceeds the threshold S_c^+

$$S^{+, \delta} = \left\{ (\xi, x) \in B_1(0) \times D; x + \delta\xi \in D \text{ and } f(r_1^+) < f(\sqrt{\delta|\xi|} S(x + \delta\xi, x, t; u)) \right\} \tag{52}$$

A straightforward calculation with $\omega(|\xi|) = 1$ shows that

$$\begin{aligned} \frac{f(r_1^+)}{\delta} \int_{S^{+, \delta}} |\xi| d\xi dx &\leq \int_{S^{+, \delta}} \frac{1}{\delta} f(\sqrt{\delta|\xi|} S(x + \delta\xi, x, t; u)) d\xi dx \\ &\leq \int_D E^\delta(x) dx = PD^\delta(u(t)) \end{aligned} \tag{53}$$

We define the weighted volume of the set $S^{+, \delta}$ to be

$$\tilde{V}(S^{+, \delta}) = \int_{S^{+, \delta}} |\xi| d\xi dx \tag{54}$$

and inequality (53) becomes

$$\frac{f(r_1^+)}{\delta} \tilde{V}(S^{+, \delta}) \leq \int_{S^{+, \delta}} \frac{1}{\delta} f(\sqrt{\delta|\xi|} S(x + \delta\xi, x, t, u)) d\xi dx \leq PD^\delta(u(t)) \tag{55}$$

Next we use Gronwall’s inequality to prove the following theorem that shows that the kinetic and peridynamic energies of the solution $u(x, t)$ are bounded by the energy put into the system.

Theorem 3.

$$C(t) \geq \frac{\rho}{2} \|\dot{u}\|_{L^2(D; \mathbb{R}^2)}^2 + PD^\delta(u(t)). \tag{56}$$

Proof. We start by multiplying both sides of (12) by \dot{u} to get

$$\rho \ddot{u}(t) \cdot \dot{u}(t) = (\mathcal{L}^T(u)(x, t) + \mathcal{L}^D(u)(x, t)) \cdot \dot{u}(t) + b(t) \cdot \dot{u}(t).$$

Applying the product rule in the first term and integration by parts in the second term gives

$$\frac{1}{2} \frac{d}{dt} \left[\rho \|\dot{u}\|_{L^2(D, \mathbb{R}^d)}^2 + 2PD^\delta(u(t)) \right] = \int_D b(t) \cdot \dot{u}(t) \, dx.$$

Application of Cauchy’s inequality to the right-hand side gives

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \left[\rho \|\dot{u}\|_{L^2(D, \mathbb{R}^d)}^2 + 2PD^\delta(u(t)) \right] &= \int_D b(t) \cdot \dot{u}(t) \, dx \\ &\leq \|b(t)\|_{L^2(D, \mathbb{R}^d)} \|\dot{u}(t)\|_{L^2(D, \mathbb{R}^d)}. \end{aligned} \tag{57}$$

Now set $\tilde{W}(t) = \rho \|\dot{u}\|_{L^2(D, \mathbb{R}^d)}^2 + 2PD^\delta(u(t)) + \zeta$ where ζ is a positive number and can be taken arbitrarily small and (57) becomes,

$$\begin{aligned} \frac{1}{2} \tilde{W}'(t) &\leq \|b(t)\|_{L^2(D, \mathbb{R}^d)} \|\dot{u}(t)\|_{L^2(D, \mathbb{R}^d)} \\ &\leq \|b(t)\|_{L^2(D, \mathbb{R}^d)} \frac{\sqrt{\tilde{W}(t)}}{\sqrt{\rho}} \end{aligned}$$

Now we can write

$$\frac{1}{2} \int_0^t \frac{\tilde{W}'(\tau)}{\sqrt{\tilde{W}(\tau)}} \, d\tau \leq \frac{1}{\sqrt{\rho}} \int_0^t \|b\|_{L^2(D, \mathbb{R}^d)} \, d\tau$$

Which simplifies to

$$\sqrt{\tilde{W}(t)} - \sqrt{\tilde{W}(0)} \leq \frac{1}{\sqrt{\rho}} \int_0^t \|b\|_{L^2(D, \mathbb{R}^d)} \, d\tau. \tag{58}$$

Since ζ can be made arbitrarily small, we find that

$$\sqrt{W(t)} - \sqrt{W(0)} \leq \frac{1}{\sqrt{\rho}} \int_0^t \|b\|_{L^2(D, \mathbb{R}^d)} \, d\tau, \tag{59}$$

and (56) follows. □

We apply inequality (55) and Theorem 3 to get the fundamental inequality.

$$\tilde{V}(S^{+, \delta}) \leq \frac{C(t)\delta}{f(\bar{r})}. \tag{60}$$

The fundamental inequality above is defined on $B_1(0) \times D$, and we now use it to bound the volume of the softening zone on D . Introducing the characteristic

function $\chi^{S^{+\delta}}(\xi, x)$ and taking the value 1 when $(\xi, x) \in S^{+\delta}$ and 0 otherwise, we immediately have

$$mP_\delta(x) = \int_{B_1(0)} \chi^{S^{+\delta}}(\xi, x) |\xi| d\xi.$$

So we can rewrite equation (54) as

$$\begin{aligned} \tilde{V}(S^{+\delta}) &= \int_D \int_{B_1(0)} \chi^{S^{+\delta}}(\xi, x) |\xi| d\xi dx \\ &= m \int_D P_\delta(x) dx. \end{aligned} \tag{61}$$

Now applying Tchebychev’s inequality to (61) with $SZ^\delta(\gamma, t)$ defined by (46) gives the desired result

$$\text{Volume}(SZ^\delta(\gamma, t)) \leq \frac{1}{\gamma} \int_D P_\delta(x) dx = \frac{\tilde{V}(S^{+\delta})}{m\gamma} \leq \frac{C(t)\delta}{m\gamma f(r_1^+)}. \tag{62}$$

Calibration of the Model

In this section we show how to calibrate this model using the known elastic properties and energy release rate of fracture associated with a given material.

Calibrating the Peridynamic Energy to Elastic Properties

We start by considering a body D for which the strain S is small. Here *small* means for a fixed $|y - x|$ we have $|S| \ll |S_c^\pm|$, $|\theta| \ll |\theta_c^\pm|$. Now we proceed to calculate the peridynamic energy density inside the material due to the presence of a small deformation $u(x)$. Suppose that the *strain* at the length scale of a neighborhood of *horizon* δ is a linear function, i.e.,

$$\begin{aligned} S(u, y, x) &= \frac{u(y) - u(x)}{|y - x|} \cdot \frac{y - x}{|y - x|} \\ &= F \frac{y - x}{|y - x|} \cdot \frac{y - x}{|y - x|} = Fe \cdot e, \end{aligned} \tag{63}$$

here F is a 3 by 3 matrix. We expand the first potential with respect to S and the second in θ keeping in mind that

$$f(0) = f'(0) = g(0) = g'(0) = 0$$

to get

$$\begin{aligned}
 f\left(\sqrt{|y-x|}S\right) &= \frac{|y-x|}{2} f''(0)S^2 + O(S^3) \\
 g\left(\theta(x,t;S)\right) &= \frac{1}{2} g''(0)\theta^2 + O(\theta^3)
 \end{aligned}
 \tag{64}$$

So we write the *energy density* which was defined in Eq. (50) for points x of distance δ away from the boundary ∂D to leading order

$$\begin{aligned}
 E^\delta &= \frac{1}{V_\delta} \frac{\alpha f''(0)}{2\delta} \int_{H_\delta(x)} \omega^\delta(|y-x|)|y-x|(Fe \cdot e)^2 dy \\
 &+ \frac{\beta g''(0)}{2\delta^2} \left(\frac{1}{V_\delta} \int_{H_\delta(x)} \omega^\delta(|y-x|)|y-x|Fe \cdot e dy \right)^2
 \end{aligned}
 \tag{65}$$

The change of variable $\delta\xi = y - x$ gives to leading order

$$\begin{aligned}
 E^\delta &= \frac{\alpha f''(0)}{2V_1} \int_{H_1(0)} \omega(|\xi|)|\xi|(Fe \cdot e)^2 d\xi \\
 &+ \frac{\beta g''(0)}{2V_1^2} \left(\int_{H_1(0)} \omega(|\xi|)|\xi|Fe \cdot e d\xi \right)^2
 \end{aligned}
 \tag{66}$$

Observe that $(Fe \cdot e)^2 = \sum_{ijkl} F_{ij} F_{kl} e_i e_j e_k e_l$ and the first term in (66) is given by

$$\sum_{ijkl} \mathbb{M}_{ijkl} F_{ij} F_{kl}
 \tag{67}$$

where

$$\begin{aligned}
 \mathbb{M}_{ijkl} &= \frac{\alpha f''(0)}{2V_1} \int_{\mathcal{H}_1(0)} |\xi| \omega(|\xi|) e_i e_j e_k e_l d\xi = \frac{\alpha f''(0)}{2V_1} \int_0^1 |\xi|^3 \omega(|\xi|) d|\xi| \\
 &\int_{S^2} e_i e_j e_k e_l de.
 \end{aligned}
 \tag{68}$$

where de is an element of surface measure on the unit sphere. Next observe $Fe \cdot e = \sum_{kj} F_{kj} e_k e_j$ and the second term in (66) is given by

$$\frac{\beta g''(0)}{2V_1^2} \left(\sum_{ij} \Lambda_{ij} F_{ij} \right)^2 = \frac{\beta g''(0)}{2V_1^2} \sum_{ijkl} \Lambda_{ij} \Lambda_{kl} F_{ij} F_{kl}.
 \tag{69}$$

where

$$\Lambda_{jk} = \int_{\mathcal{H}_1(0)} |\xi| \omega(|\xi|) e_j e_k d\xi = \int_0^1 |\xi|^3 \omega(|\xi|) d|\xi| \int_{S^2} e_j e_k de. \tag{70}$$

Focusing on the first term, we show that

$$\mathbb{M}_{ijkl} = 2\mu \left(\frac{\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}}{2} \right) + \lambda \delta_{ij} \delta_{kl} \tag{71}$$

where μ and λ are given by

$$\lambda = \mu = \frac{\alpha f''(0)}{10} \int_0^1 |\xi|^3 \omega(|\xi|) d|\xi|. \tag{72}$$

To see this we write

$$\Gamma_{ijkl}(e) = e_i e_j e_k e_l, \tag{73}$$

to observe that $\Gamma(e)$ is a totally symmetric tensor valued function defined for $e \in S^2$ with the property

$$\Gamma_{ijkl}(Qe) = Q_{im} e_m Q_{jn} e_n Q_{ko} e_o Q_{lp} e_p = Q_{im} Q_{jn} Q_{ko} Q_{lp} \Gamma_{mnop}(e) \tag{74}$$

for every rotation Q in SO^3 . Here repeated indices indicate summation. We write

$$\int_{\mathcal{H}_1(0)} |\xi|^3 \omega(|\xi|) e_i e_j e_k e_l d\xi = \int_0^1 |\xi|^3 \omega(|\xi|) d|\xi| \int_{S^2} \Gamma_{ijkl}(e) de \tag{75}$$

to see that for every Q in SO^3

$$Q_{im} Q_{jn} Q_{ko} Q_{lp} \int_{S^2} \Gamma_{ijkl}(e) de = \int_{S^2} \Gamma_{mnop}(Qe) de = \int_{S^2} \Gamma_{mnop}(e) de. \tag{76}$$

Therefore we conclude that $\int_{S^2} \Gamma_{ijkl}(e) de$ is invariant under SO^3 and is therefore an isotropic symmetric fourth-order tensor and necessarily of the form

$$\int_{S^2} \Gamma_{ijkl}(e) de = a (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}) + b \delta_{ij} \delta_{kl}. \tag{77}$$

So \mathbb{M} can be written in the form

$$\mathbb{M}_{ijkl} = 2\mu \left(\frac{\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}}{2} \right) + \lambda \delta_{ij} \delta_{kl}, \tag{78}$$

with suitable choices of μ and λ . To evaluate μ and λ , we note the following relations between μ and λ for isotropic fourth-order tensors of the form above and their contractions

$$\mathbb{M}_{ijij} = 3(2\mu + 3\lambda), \tag{79}$$

$$\mathbb{M}_{ijij} = 3(4\mu + \lambda). \tag{80}$$

These relations can be readily verified by direct calculation.

On the other hand from the definition of \mathbb{M} given by (68), we have

$$\mathbb{M}_{ijij} = \frac{\alpha f''(0)}{2V_1} \int_0^1 |\xi|^3 \omega(|\xi|) d|\xi| \int_{S^2} e_i^2 e_j^2 de = \frac{4\pi\alpha f''(0)}{2V_1} \int_0^1 |\xi|^3 \omega(|\xi|) d|\xi|, \tag{81}$$

$$\mathbb{M}_{ijij} = \frac{\alpha f''(0)}{2V_1} \int_0^1 |\xi|^3 \omega(|\xi|) d|\xi| \int_{S^2} e_i^2 e_j^2 de = \frac{4\pi\alpha f''(0)}{2V_1} \int_0^1 |\xi|^3 \omega(|\xi|) d|\xi|, \tag{82}$$

since $e_i^2 = \sum_i e_i^2 = 1$. Equation (72) now follows on recalling that $V_1 = \frac{4}{3}\pi$ and solving the system given by (79) and (80).

Focusing on the second term of (66) given by (69), we show that

$$\Lambda_{ij} = \frac{4\pi}{3} \int_0^1 |\xi|^3 \omega(|\xi|) d|\xi| \delta_{ij} \tag{83}$$

To see this we write

$$\Lambda_{ij}(e) = e_i e_j, \tag{84}$$

to observe that $\Lambda(e)$ is a totally symmetric tensor valued function defined for $e \in S^2$ with the property

$$\Lambda_{ij}(Qe) = Q_{im} e_m Q_{jn} e_n = Q_{im} Q_{jn} \Lambda_{mn}(e) \tag{85}$$

for every rotation Q in SO^3 . As before repeated indices indicate summation. We consider

$$\int_{S^2} \Lambda_{ij}(e) de \tag{86}$$

to see that for every Q in SO^3

$$Q_{im} Q_{jn} \int_{S^2} \Lambda_{ij}(e) de = \int_{S^2} \Lambda_{mn}(Qe) de = \int_{S^2} \Lambda_{mn}(e) de. \tag{87}$$

Therefore we conclude that $\int_{S^2} \Lambda_{ij}(e) de$ is an isotropic symmetric second-order tensor and of the form

$$\int_{S^2} \Lambda_{ij}(e) de = a\delta_{ij}, \tag{88}$$

i.e., a multiple of the identity. So from (70) Λ is of the form

$$\Lambda_{ij} = \gamma\delta_{ij}. \tag{89}$$

To evaluate γ we take the trace of (70) and (83) as follows.

Now the second term is given by

$$\begin{aligned} & \frac{\beta g''(0)}{2V_1^2} \left(\frac{4\pi}{3}\right)^2 \left(\int_0^1 |\xi|^3 \omega(|\xi|) d|\xi|\right)^2 \sum_{ijkl} \delta_{ij} \delta_{kl} F_{ij} F_{kl} = \\ & = \beta g''(0) \frac{1}{2} \left(\int_0^1 |\xi|^3 \omega(|\xi|) d|\xi|\right)^2 \sum_{ijkl} \delta_{ij} \delta_{kl} F_{ij} F_{kl} = \mathbb{K}_{ijkl} F_{ij} F_{kl} \end{aligned} \tag{90}$$

Collecting results we see that the leading order of the energy is given by

$$E^\delta = \sum_{ijkl} (\mathbb{M}_{ijkl} + \mathbb{K}_{ijkl}) F_{ij} F_{kl} = \sum_{ijkl} \left(2\bar{\mu} \frac{\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}}{2} + \bar{\lambda} \delta_{ij} \delta_{kl}\right) F_{ij} F_{kl} \tag{91}$$

where the shear modulus is given by

$$\bar{\mu} = \frac{\alpha f''(0)}{10} \int_0^1 |\xi|^3 \omega(|\xi|) d|\xi|, \tag{92}$$

and the Lamé constant is given by

$$\bar{\lambda} = \frac{\alpha f''(0)}{10} \int_0^1 |\xi|^3 \omega(|\xi|) d|\xi| + \frac{\beta g''(0)}{2} \left(\int_0^1 |\xi|^3 \omega(|\xi|) d|\xi|\right)^2. \tag{93}$$

One is free to choose α and β provided that the resulting elastic tensor satisfies the constraints of ellipticity. Here one is no longer restricted to Poisson ratios of 1/4 as in the bond-based formulation.

An identical calculation shows that for two-dimensional problems the elastic constants are given by

$$\bar{\mu} = \frac{\alpha f''(0)}{8} \int_0^1 |\xi|^2 \omega(|\xi|) d|\xi|, \tag{94}$$

and

$$\bar{\lambda} = \frac{\alpha f''(0)}{8} \int_0^1 |\xi|^2 \omega(|\xi|) d|\xi| + \frac{\beta g''(0)}{2} \left(\int_0^1 |\xi|^2 \omega(|\xi|) d|\xi| \right)^2, \quad (95)$$

and one is no longer restricted to Poisson ratio 1/3 materials.

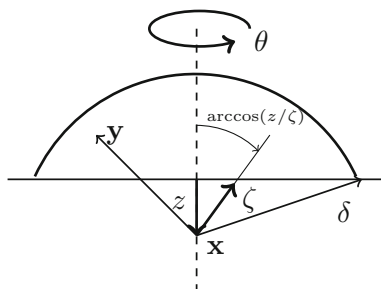
We note here that the two-dimensional moduli $\bar{\mu}$ and $\bar{\lambda}$ are directly related to the well-known moduli appearing in the plane strain or plane stress solutions for isotropic materials. This relationship is now well known and can be found in Jasiuk et al. (1994) and also Milton (2002).

Calibrating Energy Release Rate

In regions of large strain, the same force potentials (3) and (4) are used to calculate the amount of energy consumed by a crack per unit area of growth, i.e., the energy release rate. The energy release rate equals the work necessary to eliminate force interaction on either side of a fracture surface per unit fracture area. In this model the energy release rate has two components: one associated with the force potential for tensile strain (3) and the other associated with the force potential for hydrostatic strain (4). The critical energy release rate \mathcal{G}_s associated with fracture under tensile forces is found to be the same for all choices of horizon δ . However the critical energy release rate for hydrostatic fracture \mathcal{G}_h increases with decreasing horizon and becomes infinite as $\delta \rightarrow 0$ at the rate $1/\delta$.

For tensile forces we use (3) and calculate the work required to eliminate interaction between two points x and y ; this is given by $\mathcal{W}^\delta(\infty) = \lim_{S \rightarrow \infty} \mathcal{W}^\delta(S)$ where $\mathcal{W}^\delta(\infty) = \omega^\delta(|y - x|) f_\infty / \delta$. We suppose x gives the center of the peridynamic neighborhood located a distance z away from the planar interface separating upper and lower half spaces. We suppose x lies in the lower half space, and the points y lie in the upper half space inside the peridynamic neighborhood of x (see Fig. 4). The critical energy release rate \mathcal{G}_s associated with tensile forces equals the work necessary to eliminate force interaction on either side of a fracture surface per unit fracture area. It is given in three dimensions by integration of $\mathcal{W}^\delta(\infty)$ over

Fig. 4 Evaluation of energy release rate \mathcal{G}_s . For each point x along the dashed line, $0 \leq z \leq \delta$, the work required to break the interaction between x and y in the spherical cap is summed up in (96) using spherical coordinates centered at x



the intersection of the neighborhood of x and the upper half space given by the spherical cap (see Fig. 4),

$$\mathcal{G}_s = \frac{4\pi}{V_\delta} \int_0^\delta \int_z^\delta \int_0^{\cos^{-1}(z/\zeta)} \mathcal{W}^\delta(\infty, \zeta) \zeta^2 \sin \phi \, d\phi \, d\zeta \, dz \tag{96}$$

where $\zeta = |y - x|$. This integral is calculated and for d dimensions $d = 1, 2, 3$, the result is

$$\mathcal{G}_s = M \frac{2\omega_{d-1}}{\omega_d} f_\infty, \tag{97}$$

where $M = \int_0^1 r^d \omega(r) dr$ and ω_d is the volume of the d dimensional unit ball, $\omega_1 = 2, \omega_2 = \pi, \omega_3 = 4\pi/3$. We see from this calculation that the critical energy release rate is independent of δ .

For hydrostatic forces we use (4) and calculate the work required to eliminate interaction between x and the upper half plane. As before we suppose x gives the center of the peridynamic neighborhood located a distance z away from the planar interface separating upper and lower half spaces. We suppose x lies in the lower half space, and the peridynamic neighborhood of x intersects the upper half space (see Fig. 5).

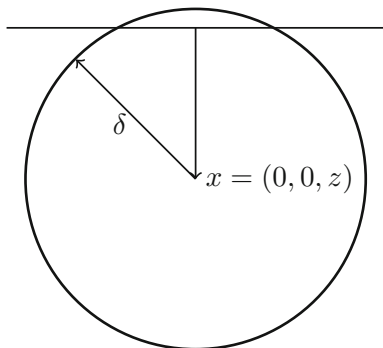
The critical energy release rate \mathcal{G}_h associated with hydrostatic forces equals the work necessary to eliminate force interaction on either side of a fracture surface per unit fracture area. The work per unit volume needed to eliminate hydrostatic interaction between a point x and its neighbors is

$$\mathcal{V}^\delta(\infty)(x) = \lim_{\theta \rightarrow \infty} \frac{\beta g(\theta)}{\delta^2} = \frac{\beta g_\infty}{\delta^2}. \tag{98}$$

For points $x = (0, 0, z)$, with $0 < |z| < \delta$ above and below the $z = 0$ plane, the work per unit area to eliminate hydrostatic interaction between the lower half space $z < 0$ and upper half space $z > 0$ is

$$\mathcal{G}_h = 2 \int_0^\delta \frac{\beta g_\infty}{\delta^2} dz = \frac{2\beta g_\infty}{\delta}. \tag{99}$$

Fig. 5 Hydrostatic energy release rate \mathcal{G}_h



For d dimensions $d = 1, 2, 3$, the result is the same and

$$G_s = \frac{2\beta g_\infty}{\delta}. \tag{100}$$

We see from this calculation that the energy release rate for hydrostatic fracture is increasing at the rate $1/\delta$.

Linear Elastic Operators in the Limit of Vanishing Horizon

In this section we consider smooth evolutions u in space and show that away from fracture set the operators $\mathcal{L}^T + \mathcal{L}^D$ acting on u converge to the operator of linear elasticity in the limit of vanishing nonlocality. We denote the fracture set by \tilde{D} and consider any open un-fractured set D' interior to D with its boundary a finite distance away from the boundary of D and the fracture set \tilde{D} . In what follows we suppose that the nonlocal horizon δ is smaller than the distance separating the boundary of D' from the boundaries of D and \tilde{D} .

Theorem 4. *Convergence to linear elastic operators. Suppose that $u(x, t) \in C^2([0, T_0], C^3(D, \mathbb{R}^3))$ and for every $x \in D' \subset D \setminus \tilde{D}$, then there is a constant $C > 0$ independent of nonlocal horizon δ such that, for every (x, t) in $D' \times [0, T_0]$, one has*

$$|\mathcal{L}^T(u(t)) + \mathcal{L}^D(u(t)) - \nabla \cdot \mathbb{C} \mathcal{E}(u(t))| < C\delta, \tag{101}$$

where the elastic strain is $\mathcal{E}(u) = (\nabla u + (\nabla u)^T)/2$ and the elastic tensor is isotropic and given by

$$\mathbb{C}_{ijkl} = 2\bar{\mu} \left(\frac{\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}}{2} \right) + \bar{\lambda}\delta_{ij}\delta_{kl}, \tag{102}$$

with shear modulus $\bar{\mu}$ and Lamé coefficient $\bar{\lambda}$ given by (92) and (93). The numbers α and β can be chosen independently and can be any pair of real numbers such that \mathbb{C} is positive definite.

Proof. We start by showing

$$|\mathcal{L}^T(u(t)) - \frac{f''(0)}{2\omega_3} \int_{B_1(0)} e|\xi|J(|\xi|)e_i e_j e_k d\xi \partial_{jk}^2 u_i(x)| < C\delta, \tag{103}$$

where $\omega_3 = 4\pi/3$ and $e = e_{y-x}$ are unit vectors on the sphere; here repeated indices indicate summation. To see this recall the formula for $\mathcal{L}^T(u)$ and write $\partial_S f(\sqrt{|y-x|S}) = f'(\sqrt{|y-x|S})\sqrt{|y-x|}$. Now Taylor expand

$f'(\sqrt{|y-x|}S)$ in $\sqrt{|y-x|}S$ and Taylor expand $u(y)$ about x , denoting e_{y-x} by e to find that all odd terms in e integrate to zero and

$$|\mathcal{L}^T(u(t))_l - \frac{2}{V_\delta} \int_{B_\delta(x)} \frac{J^\delta(|y-x|)}{\delta|y-x|} \frac{f''(0)}{4} |y-x|^2 \partial_{jk}^2 u_i(x) e_i e_j e_k e_l, dy| < C\delta, l = 1, 2, 3. \tag{104}$$

On changing variables $\xi = (y-x)/\delta$, we recover (103). Now we show

$$|\mathcal{L}^D(u(t))_k - \frac{1}{\omega_3} \int_{B_1(0)} |\xi| \omega(|\xi|) e_i e_j d\xi \frac{\beta g''(0)}{2\omega_3} \int_{B_1(0)} |\xi| \omega(|\xi|) e_k e_l d\xi \partial_{lj}^2 u_i(x)| < C\delta, k = 1, 2, 3. \tag{105}$$

We note for $x \in D'$ that $D \cap B_\delta(x) = B_\delta(x)$ and the integrand in the second term of (14) is odd and the integral vanishes. For the first term in (14), we Taylor expand $\partial_\theta g(\theta)$ about $\theta = 0$ and Taylor expand $u(z)$ about y inside $\theta(y, t)$ noting that terms odd in $e = e_{z-y}$ integrate to zero to get

$$|\partial_\theta g(\theta(y, t)) - g''(0) \frac{1}{V_\delta} \int_{B_\delta(y)} \omega^\delta(|z-y|) |z-y| \partial_j u_i(y) e_i e_j dz| < C\delta^3. \tag{106}$$

Now substitution for the approximation to $\partial_\theta g(\theta(y, t))$ in the definition of \mathcal{L}^D gives

$$\left| \mathcal{L}^D(u) \frac{1}{V_\delta} \int_{B_\delta(x)} \frac{\omega^\delta(|y-x|)}{\delta^2} e_{y-x} \frac{1}{2V_\delta} \int_{B_\delta(y)} \omega^\delta(|z-y|) |z-y| \beta g''(0) \partial_j u_i(y) e_i e_j dz dy \right| < C\delta. \tag{107}$$

We Taylor expand $\partial_j u_i(y)$ about x ; note that odd terms involving tensor products of e_{y-x} vanish when integrated with respect to y in $B_\delta(x)$, and we obtain (105).

We now calculate as in (Lipton 2016 equation (64) or in section “Calibrating the Peridynamic Energy to Elastic Properties” to find that

$$\begin{aligned} & \frac{f''(0)}{2\omega_3} \int_{B_1(0)} |\xi| J(|\xi|) e_i e_j e_k e_l d\xi \partial_{jk}^2 u_i(x) \\ &= \left(2\mu_1 \left(\frac{\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}}{2} \right) + \lambda_1 \delta_{ij} \delta_{kl} \right) \partial_{jk}^2 u_i(x), \end{aligned} \tag{108}$$

where

$$\mu_1 = \lambda_1 = \frac{f''(0)}{10} \int_0^1 r^3 \omega(r) dr. \tag{109}$$

Next observe that a straightforward calculation gives

$$\frac{1}{\omega_3} \int_{B_1(0)} |\xi| \omega(|\xi|) e_i e_j d\xi = \delta_{ij} \int_0^1 r^3 \omega(r) dr, \quad (110)$$

and we deduce that

$$\begin{aligned} & \frac{1}{\omega_3} \int_{B_1(0)} |\xi| \omega(|\xi|) e_i e_j d\xi \frac{\beta g''(0)}{2\omega_3} \int_{B_1(0)} |\xi| \omega(|\xi|) e_k e_l d\xi \partial_{ij}^2 u_i(x) \\ &= \frac{\beta g''(0)}{2} \left(\int_0^1 r^3 \omega(r) dr \right)^2 \delta_{ij} \delta_{kl} \partial_{ij}^2 u_i(x). \end{aligned} \quad (111)$$

Theorem 4 follows on adding (108) and (111) □

Conclusions

We have introduced a regularized model for free fracture propagation based on nonlocal potentials. At each instant of the evolution, we identify the softening zone where strains lie above the strength of the material. We have shown that discontinuities associated with flaws larger than the length scale of nonlocality δ can become unstable and grow. An explicit inequality is found that shows that the volume of the softening zone goes to zero linearly with the length scale of nonlocal interaction. This scaling is consistent with the notion that a softening zone of width proportional to δ converges to a sharp fracture as the length scale of nonlocal interaction goes to zero. Inside quiescent regions with no cracks, the nonlocal operator converges to the local elastic operator at a rate proportional to the radius of nonlocal interaction. We show that the model can be calibrated to measured values of critical energy release rate, shear modulus, and bulk modulus of material samples. The double-well state-based potential developed here no longer has Poisson ratio constrained to be $1/4$. For this model we can choose the potentials so that the small strain behavior is specified by the isotropic elasticity tensor for any material with prescribed shear and Lamé moduli.

The energy release rate necessary for tensile forces to create fractures is constant in δ , whereas the forces necessary to create a fracture using hydrostatic forces grows as $1/\delta$. Thus creation of fracture surfaces by hydrostatic forces will not be seen when

$$\delta < \frac{2\beta g_\infty}{G_s}. \quad (112)$$

On the other hand, the elastic properties for small strains can be made to correspond to any positive definite isotropic elastic tensor.

Acknowledgements This material is based upon the work supported by the US Army Research Laboratory and the US Army Research Office under contract/grant number W911NF1610456.

References

- A. Agwai, I. Guven, E. Madenci, Predicting crack propagation with peridynamics: a comparative study. *Int. J. Fract.* **171**, 65–78 (2011)
- F. Bobaru, W. Hu, The meaning, selection, and use of the peridynamic horizon and its relation to crack branching in brittle materials. *Int. J. Fract.* **176**, 215–222 (2012)
- F. Bobaru, J.T. Foster, P.H. Geubelle, S.A. Silling, *Handbook of Peridynamic Modeling* (CRC Press, Boca Raton, 2016)
- K. Dayal, K. Bhattacharya, Kinetics of phase transformations in the peridynamic formulation of continuum mechanics. *J. Mech. Phys. Solids* **54**, 1811–1842 (2006)
- P. Diehl, R. Lipton, M.A. Schweitzer, Numerical verification of a bond-based softening peridynamic model for small displacements: deducing material parameters from classical linear theory. Institut für Numerische Simulation Preprint No. 1630 (2016)
- B.K. Driver, *Analysis Tools with Applications*. E-book (Springer, Berlin, 2003)
- Q. Du, M. Gunzburger, R. Lehoucq, K. Zhou, Analysis of the volume-constrained peridynamic Navier equation of linear elasticity. *J. Elast.* **113**, 193–217 (2013)
- J.T. Foster, S.A. Silling, W. Chen, An energy based failure criterion for use with peridynamic states. *Int. J. Multiscale Comput. Eng.* **9**, 675–688 (2011)
- W. Gerstle, N. Sau, S. Silling, Peridynamic modeling of concrete structures. *Nucl. Eng. Des.* **237**, 1250–1258 (2007)
- Y.D. Ha, F. Bobaru, Studies of dynamic crack propagation and crack branching with peridynamics. *Int. J. Fract.* **162**, 229–244 (2010)
- I. Jasiuk, J. Chen, M.F. Thorpe, Elastic moduli of two dimensional materials with polygonal and elliptical holes. *Appl. Mech. Rev.* **47**, S18–S28 (1994)
- R. Lipton, Dynamic brittle fracture as a small horizon limit of peridynamics. *J. Elast.* **117**, 21–50 (2014)
- R. Lipton, Cohesive dynamics and brittle fracture. *J. Elast.* **124**(2), 143–191 (2016)
- R. Lipton, S. Silling, R. Lehoucq, Complex fracture nucleation and evolution with nonlocal elastodynamics. arXiv preprint arXiv:1602.00247 (2016)
- G.W. Milton, *The Theory of Composites* (Cambridge University Press, Cambridge, 2002)
- S.A. Silling, Reformulation of elasticity theory for discontinuities and long-range forces. *J. Mech. Phys. Solids* **48**, 175–209 (2000)
- S.A. Silling, E. Askari, A meshfree method based on the peridynamic model of solid mechanics. *Comput. Struct* **83**, 1526–1535 (2005)
- S.A. Silling, F. Bobaru, Peridynamic modeling of membranes and fibers. *Int. J. Nonlinear Mech.* **40**, 395–409 (2005)
- S.A. Silling, R.B. Lehoucq, Convergence of peridynamics to classical elasticity theory. *J. Elast.* **93**, 13–37 (2008)
- S. A. Silling, M. Epton, O. Weckner, J. Xu, E. Askari, Peridynamic states and constitutive modeling. *J. Elast.* **88**, 151–184 (2007)
- S. Silling, O. Weckner, E. Askari, F. Bobaru, Crack nucleation in a peridynamic solid. *Int. J. Fract.* **162**, 219–227 (2010)
- O. Weckner, R. Abeyaratne, The effect of long-range forces on the dynamics of a bar. *J. Mech. Phys. Solids* **53**, 705–728 (2005)



Nonlocal Operators with Local Boundary Conditions: An Overview

39

Burak Aksoylu, Fatih Celiker, and Orsan Kilicer

Contents

Introduction	1294
The Novel Operators in 2D	1296
Operators in 1D	1302
The Construction of 2D Operators	1311
Verifying the Boundary Conditions	1314
Operators in Higher Dimensions	1316
Numerical Experiments	1317
The Treatment of General Nonlocal Problems Using Functional Calculus	1318
Conclusion	1328
References	1329

Abstract

We present novel governing operators in arbitrary dimension for nonlocal diffusion in homogeneous media. The operators are inspired by the theory of peridynamics (PD). They agree with the original PD operator in the bulk of the domain and simultaneously enforce local boundary conditions (BC). The main ingredients are periodic, antiperiodic, and mixed extensions of kernel functions together with even and odd parts of bivariate functions. We present different types of BC in 2D which include pure and mixed combinations of Neumann and Dirichlet BC. Our construction is systematic and easy to follow. We provide numerical experiments that validate our theoretical findings. When our novel operators are extended to vector-valued functions, they will allow the

Mathematics Subject Classification (2000): 35L05, 74B99, 47G10.

B. Aksoylu (✉) · F. Celiker · O. Kilicer

Department of Mathematics, Wayne State University, Detroit, MI, USA

e-mail: burak@wayne.edu; celiker@wayne.edu; okilicer@math.tamu.edu

© Springer Nature Switzerland AG 2019

G. Z. Voyiadjis (ed.), *Handbook of Nonlocal Continuum Mechanics for Materials and Structures*, https://doi.org/10.1007/978-3-319-58729-5_34

1293

extension of PD to applications that require local BC. Furthermore, we hope that the ability to enforce local BC provides a remedy for surface effects seen in PD.

We recently proved that the nonlocal diffusion operator is a function of the classical operator. This observation opened a gateway to incorporate local BC to nonlocal problems on bounded domains. The main tool we use to define the novel governing operators is functional calculus, in which we replace the classical governing operator by a suitable function of it. We present how to apply functional calculus to general nonlocal problems in a methodical way.

Keywords

Nonlocal wave equation · Nonlocal operator · Peridynamics · Boundary condition · Integral operator

Introduction

We construct novel governing operators for nonlocal diffusion (Andreu-Vaillio et al. 2010; Du et al. 2012) in arbitrary dimension. The operators are inspired by the theory of peridynamics (PD), a nonlocal extension of continuum mechanics developed by Silling (2000). By suppressing the time variable t , we take the following operator as the original governing operator, and, in 1D, it corresponds to the original bond-based PD operator for homogeneous media. We choose the 2D domain as $\Omega := [-1, 1] \times [-1, 1]$ and for $(x, y) \in \Omega$,

$$\begin{aligned} \mathcal{L}_{\text{orig}}u(x, y) := & \iint_{\Omega} \widehat{C}(x' - x, y' - y)u(x, y)dx'dy' \\ & - \iint_{\Omega} \widehat{C}(x' - x, y' - y)u(x', y')dx'dy'. \end{aligned} \quad (1)$$

We define the operator that is closely related to $\mathcal{L}_{\text{orig}}$ as

$$\mathcal{L}u(x, y) := cu(x, y) - \iint_{\Omega} \widehat{C}(x' - x, y' - y)u(x', y')dx'dy', \quad (x, y) \in \Omega. \quad (2)$$

where $C := \widehat{C}|_{\Omega}$ and $c := \iint_{\Omega} C(x', y')dx'dy'$. Since \widehat{C} enters into the formulation only as a function of $(x' - x, y' - y)$, the operator \mathcal{L} also assumes a homogeneous medium. We will show that the two operators agree in the bulk. As the main contribution, we prove that the novel governing operators we construct agree with

$\mathcal{L}_{\text{orig}}$ in the bulk of Ω and, at the same time, enforce local pure and mixed Neumann, Dirichlet, periodic, and antiperiodic BC.

When PD is considered, the dimension of u must be equal to that of x . In that case, the governing operator in (1) restricted to 1D corresponds to the bond-based linearized PD; see Silling et al. (2003, Eq. 23) and Weckner and Abeyaratne (2005, Eq. 3). For the discussion of PD, it is implied that $u, x \in \mathbb{R}$. The case of $u \in \mathbb{R}$ and $x \in \mathbb{R}^d$ corresponds to nonlocal diffusion (Du et al. 2012; Seleson et al. 2013).

Our approach to nonlocal problems is fundamentally different because we exclusively want to use local BC. In Beyer et al. (2016), one of our major results was the finding that the governing operator of PD equation in \mathbb{R} and nonlocal diffusion in \mathbb{R}^d are functions of the Laplace operator. This result opened the path to the introduction of local boundary conditions into PD theory. Since PD is a nonlocal theory, one might expect only the appearance of nonlocal BC while employing $\mathcal{L}_{\text{orig}}$ as the governing operator. In the original PD formulation, the concept of local BC does not apply to PD. Instead, external forces must be supplied through the loading force density (Silling 2000, p.201). On the other hand, we demonstrate that the anticipation that local BC are incompatible with nonlocal operators is not quite correct. Our novel operators present an alternative to nonlocal BC, and we hope that the ability to enforce local BC will provide a remedy for surface effects seen in PD; see Madenci and Oterkus (2014, Chaps. 4, 5, 7, and 12) and Kilic (2008), Mitchell et al. (2015). Furthermore, our approach will provide us the capability to solve important elasticity problems that require local BC such as contact, shear, and traction.

We studied various aspects of local BC in nonlocal problems (Aksoylu et al. 2017a,b, 2016, 2017, Submitted; Beyer et al. 2016). Building on Beyer et al. (2016), we generalized the results in \mathbb{R} to bounded domains (Aksoylu et al. 2017a,b), a critical feature for all practical applications. In Aksoylu et al. (2017b), we laid the theoretical foundations, and in Aksoylu et al. (2017a), we applied the foundations to prominent BC such as Dirichlet and Neumann, as well as presented numerical implementation of the corresponding wave propagation. In Aksoylu et al. (2017), we constructed the first 1D operators that agree with the original bond-based PD operator in the bulk of the domain and simultaneously enforce local Neumann and Dirichlet BC which we denote by \mathcal{M}_{N} and \mathcal{M}_{D} , respectively. We carried out numerical experiments by utilizing \mathcal{M}_{N} and \mathcal{M}_{D} as governing operators in Aksoylu et al. (2017a). In Aksoylu et al. (2016), we studied other related governing operators. In Aksoylu and Kaya (2018), we study the condition numbers of the novel governing operators. Therein, we prove that the modifications made to the operator $\mathcal{L}_{\text{orig}}$ to obtain the novel operators are minor as far as the condition numbers are concerned.

Our approach is not limited to PD; the abstractness of the theoretical methods used allows generalization to other nonlocal theories. Our approach presents a unique way of combining the powers of abstract operator theory with numerical computing (Aksoylu et al. 2017a). Similar classes of operators are used in numerous applications such as nonlocal diffusion (Andreu-Vaillo et al. 2010; Du et al. 2012; Seleson et al. 2013), image processing (Gilboa and Osher 2008), population models, particle systems, phase transition, and coagulation. See the review and news articles

Du et al. (2012, 2014), and Silling and Lehoucq (2010) for a comprehensive discussion and the book (Madenci and Oterkus 2014). In addition, see the studies dedicated to conditioning analysis, domain decomposition and variational theory (Aksoylu and Kaya 2018; Aksoylu and Mengesha 2010; Aksoylu and Parks 2011; Aksoylu and Unlu 2014), discretization (Aksoylu and Unlu 2014; Emmrich and Weckner 2007; Tian and Du 2013), and kernel functions (Mengesha and Du 2013; Seleson and Parks 2011).

The rest of the paper is structured as follows. In section “[The Novel Operators in 2D](#),” first we prove that the operators $\mathcal{L}_{\text{orig}}$ and \mathcal{L} agree in the bulk in 2D. We define the novel operators using orthogonal projections on bivariate functions for which we utilize the periodic, antiperiodic, and mixed extensions of the kernel function $C(x)$. We give the main theorem in 2D. In section “[Operators in 1D](#),” first we prove that the novel operators are self-adjoint. In 1D, we give the main theorem which states they all agree with $\mathcal{L}_{\text{orig}}$ in the bulk and simultaneously enforce the corresponding local BC. In section “[The Construction of 2D Operators](#),” we exploit the properties of the operators in 1D to construct the novel operators in 2D. We transfer the agreement in the bulk property established for univariate functions to bivariate ones and eventually prove that the novel operators agree with $\mathcal{L}_{\text{orig}}$ in the bulk in 2D. In section “[Verifying the Boundary Conditions](#),” we make use of the Leibniz rule, the Fubini theorem, and the Lebesgue dominated convergence theorem to prove that the novel operators enforce the local BC stated in the main theorem. In section “[Operators in Higher Dimensions](#),” we present the operators in 3D which can be easily extended to arbitrary dimension. In section “[Numerical Experiments](#),” we report the numerical experiments. In section “[The Treatment of General Nonlocal Problems Using Functional Calculus](#),” we present the treatment of general nonlocal problems using functional calculus. We conclude in section “[Conclusion](#).”

The Novel Operators in 2D

For $(x, y), (x', y') \in \Omega^1$, it follows that $(x' - x, y' - y) \in [-2, 2] \times [-2, 2]$. Hence, in (1), the domain of $\widehat{C}(x' - x, y' - y)$ is $\widehat{\Omega} := [-2, 2] \times [-2, 2]$. Furthermore, the kernel function $\widehat{C}(x, y)$ is assumed to be even. Namely,

$$\widehat{C}(-x, -y) = \widehat{C}(x, y).$$

The important choice of $\widehat{C}(x, y)$ is the *canonical* kernel function $\widehat{\chi}_\delta(x, y)$ whose only role is the representation of the nonlocal neighborhood, called the *horizon*, by a characteristic function. More precisely, for $(x, y) \in \widehat{\Omega}$,

¹We do not explicitly denote the dimension of the domain Ω . The dimension is implied by the number of iterated integrals present in the operator. The domain Ω is equal to $[-1, 1]$, $[-1, 1] \times [-1, 1]$, and $[-1, 1] \times [-1, 1] \times [-1, 1]$ in 1D, 2D, and 3D, respectively.

$$\widehat{\chi}_\delta(x, y) := \begin{cases} 1, & (x, y) \in (-\delta, \delta) \times (-\delta, \delta) \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

The size of nonlocality is determined by δ and we assume $\delta < 1$. Since the horizon is constructed by $\widehat{\chi}_\delta(x, y)$, a kernel function used in practice is in the form

$$\widehat{C}(x, y) = \widehat{\chi}_\delta(x, y)\widehat{\mu}(x, y), \quad (4)$$

where $\widehat{\mu}(x, y) \in L^2(\widehat{\Omega})$ is even.

Throughout the paper, we assume that

$$u(x, y) \in L^2(\Omega) \cap C^1(\partial\Omega). \quad (5)$$

Inspired by the projections that give the even and odd parts of a univariate function, we define the following operators that act on a bivariate function:

$$P_{e,x'}, P_{o,x'}, P_{e,y'}, P_{o,y'} : L^2(\Omega) \rightarrow L^2(\Omega),$$

whose definitions are

$$P_{e,x'}u(x', y') := \frac{u(x', y') + u(-x', y')}{2}, \quad P_{o,x'}u(x', y') := \frac{u(x', y') - u(-x', y')}{2}, \quad (6)$$

$$P_{e,y'}u(x', y') := \frac{u(x', y') + u(x', -y')}{2}, \quad P_{o,y'}u(x', y') := \frac{u(x', y') - u(x', -y')}{2}. \quad (7)$$

Each operator is an orthogonal projection and possesses the following decomposition property:

$$P_{e,x'} + P_{o,x'} = I_{x'}, \quad P_{e,y'} + P_{o,y'} = I_{y'}. \quad (8)$$

One can easily check that all four orthogonal projections in (6) and (7) commute with each other. We define the following new operators obtained from the products of these projections:

$$P_{e,x'}P_{e,y'}u(x', y') := \frac{1}{4}\{[u(x', y') + u(x', -y')] + [u(-x', y') + u(-x', -y')]\},$$

$$P_{e,x'}P_{o,y'}u(x', y') := \frac{1}{4}\{[u(x', y') - u(x', -y')] + [u(-x', y') - u(-x', -y')]\},$$

$$P_{o,x'}P_{o,y'}u(x', y') := \frac{1}{4}\{[u(x', y') - u(x', -y')] - [u(-x', y') - u(-x', -y')]\},$$

$$P_{o,x'}P_{e,y'}u(x', y') := \frac{1}{4}\{[u(x', y') + u(x', -y')] - [u(-x', y') + u(-x', -y')]\}.$$

Due to the aforementioned commutativity property, these are also orthogonal projections and satisfy the following decomposition property:

$$P_{e,x'}P_{e,y'} + P_{e,x'}P_{o,y'} + P_{o,x'}P_{e,y'} + P_{o,x'}P_{o,y'} = I_{x',y'}. \tag{9}$$

These will be used in the definition of the novel operators in 2D.

Reflecting on the square support of the restricted kernel function $\chi_\delta(x, y)$, we define the bulk of the domain as follows:

$$\text{bulk} = \{(x, y) \in \Omega : (x, y) \in (-1 + \delta, 1 - \delta) \times (-1 + \delta, 1 - \delta)\}.$$

We first prove that the operators \mathcal{L} and $\mathcal{L}_{\text{orig}}$ agree in the bulk. Throughout the paper, we denote the restriction of a function $\widehat{Z} : \widehat{\Omega} \rightarrow \mathbb{R}$ to Ω as Z , i.e., $Z := \widehat{Z}|_\Omega$.

Lemma 1.

$$\mathcal{L}u(x, y) = \mathcal{L}_{\text{orig}}u(x, y), \quad (x, y) \in \text{bulk}.$$

Proof. For $(x, y) \in \text{bulk}$, we have

$$(x - \delta, x + \delta) \times (y - \delta, y + \delta) \cap \Omega = (x - \delta, x + \delta) \times (y - \delta, y + \delta).$$

Hence,

$$\begin{aligned} \iint_{\Omega} \widehat{C}(x' - x, y' - y) dx' dy' &= \iint_{\Omega} \widehat{\chi}_\delta(x' - x, y' - y) \widehat{\mu}(x' - x, y' - y) dx' dy' \\ &= \int_{x-\delta}^{x+\delta} \int_{y-\delta}^{y+\delta} \widehat{\mu}(x' - x, y' - y) dx' dy' = \int_{-\delta}^{\delta} \int_{-\delta}^{\delta} \mu(x', y') dx' dy' \\ &= \iint_{\Omega} \chi_\delta(x', y') \mu(x', y') dx' dy' = \iint_{\Omega} C(x', y') dx' dy'. \end{aligned}$$

The result follows. □

In the construction of the novel operators, a crucial ingredient is first restricting \widehat{C} to Ω and then suitably extending it back to $\widehat{\Omega}$. To this end, we define the periodic, antiperiodic, and mixed extensions of $C(x)$ from $[-1, 1]$ to $[-2, 2]$, respectively, as follows:

$$\widehat{C}_p(x) := \begin{cases} C(x+2), & x \in [-2, -1), \\ C(x), & x \in [-1, 1], \\ C(x-2), & x \in (1, 2], \end{cases} \quad \widehat{C}_a(x) := \begin{cases} -C(x+2), & x \in [-2, -1), \\ C(x), & x \in [-1, 1], \\ -C(x-2), & x \in (1, 2]. \end{cases} \tag{10}$$

We also utilize the following mixed extensions of $C(x)$:

$$\widehat{C}_{pa}(x) := \begin{cases} C(x+2), & x \in [-2, -1), \\ C(x), & x \in [-1, 1], \\ -C(x-2), & x \in (1, 2], \end{cases} \quad \widehat{C}_{ap}(x) := \begin{cases} -C(x+2), & x \in [-2, -1), \\ C(x), & x \in [-1, 1], \\ C(x-2), & x \in (1, 2]. \end{cases}$$

Building on our 1D construction in Aksoylu et al. (2017), in higher dimensions, we discovered the operators that enforce local pure and mixed Neumann and Dirichlet BC. We present the main theorem in 2D with the following 4 types of BC.

Theorem 1 (Main Theorem in 2D). *Let $\Omega := [-1, 1] \times [-1, 1]$ and the restricted kernel function be separable in the form*

$$C(x, y) = X(x)Y(y), \tag{11}$$

where X and Y are even functions. Then, the operators \mathcal{M}_N , \mathcal{M}_D , $\mathcal{M}_{ND,ND}$, and $\mathcal{M}_{N,DN}$ defined by

$$\begin{aligned} (\mathcal{M}_N - c)u(x, y) &:= - \iint_{\Omega} [\widehat{X}_p(x' - x)P_{e,x'} + \widehat{X}_a(x' - x)P_{o,x'}] \\ &\quad [\widehat{Y}_p(y' - y)P_{e,y'} + \widehat{Y}_a(y' - y)P_{o,y'}]u(x', y')dx' dy', \\ (\mathcal{M}_D - c)u(x, y) &:= - \iint_{\Omega} [\widehat{X}_a(x' - x)P_{e,x'} + \widehat{X}_p(x' - x)P_{o,x'}] \\ &\quad [\widehat{Y}_a(y' - y)P_{e,y'} + \widehat{Y}_p(y' - y)P_{o,y'}]u(x', y')dx' dy', \\ (\mathcal{M}_{ND,ND} - c)u(x, y) &:= - \iint_{\Omega} [\widehat{X}_{ap}(x' - x)P_{e,x'} + \widehat{X}_{pa}(x' - x)P_{o,x'}] \\ &\quad [\widehat{Y}_{ap}(y' - y)P_{e,y'} + \widehat{Y}_{pa}(y' - y)P_{o,y'}]u(x', y')dx' dy', \\ (\mathcal{M}_{N,DN} - c)u(x, y) &:= - \iint_{\Omega} [\widehat{X}_p(x' - x)P_{e,x'} + \widehat{X}_a(x' - x)P_{o,x'}] \\ &\quad [\widehat{Y}_{pa}(y' - y)P_{e,y'} + \widehat{Y}_{ap}(y' - y)P_{o,y'}]u(x', y')dx' dy', \end{aligned}$$

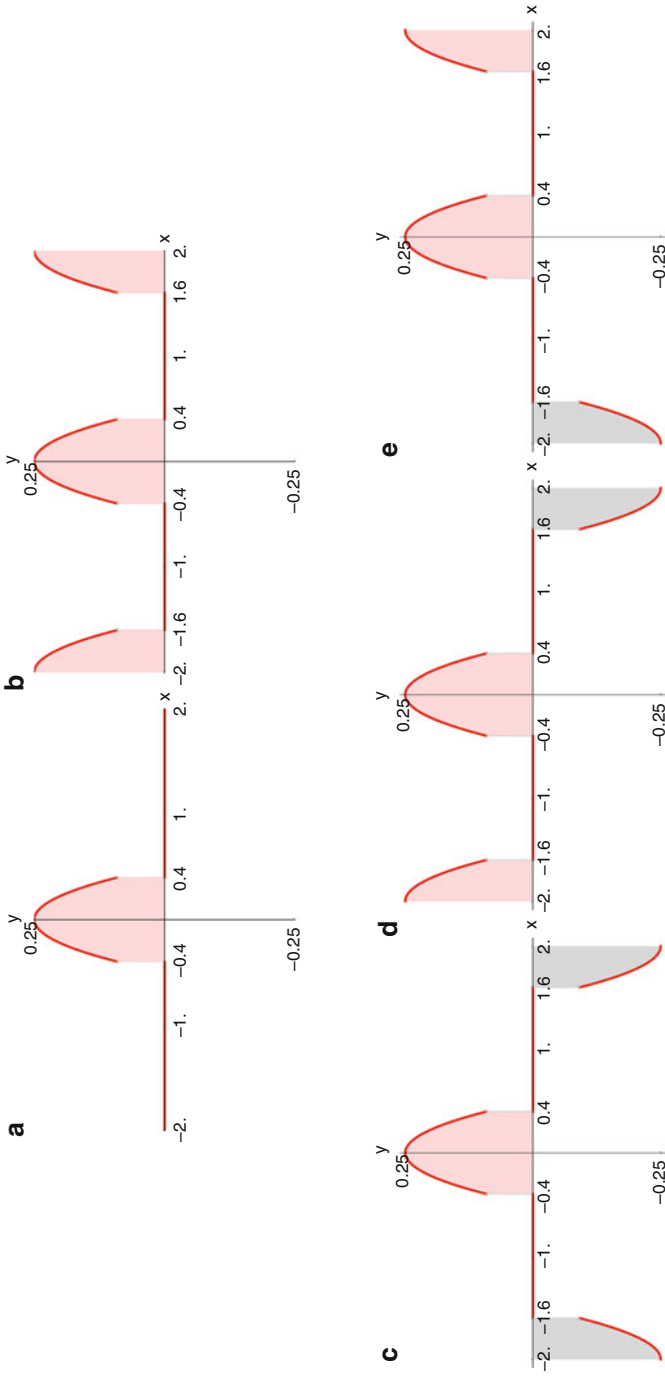


Fig. 1 The kernel function $\widehat{C}(x) = \widehat{\chi}_\delta(x)\widehat{\mu}(x)$ with $\delta = 0.4$ and $\widehat{\mu}(x) = 0.25 - x^2$. The periodic, antiperiodic, and mixed extensions of $C = \widehat{C}|_\Omega$ are denoted by \widehat{C}_p , \widehat{C}_a , \widehat{C}_{pa} , and \widehat{C}_{ap} , respectively. Plots of (a) $\widehat{C}(x)$, (b) $\widehat{C}_p(x)$, (c) $\widehat{C}_a(x)$, (d) $\widehat{C}_{pa}(x)$, (e) $\widehat{C}_{ap}(x)$

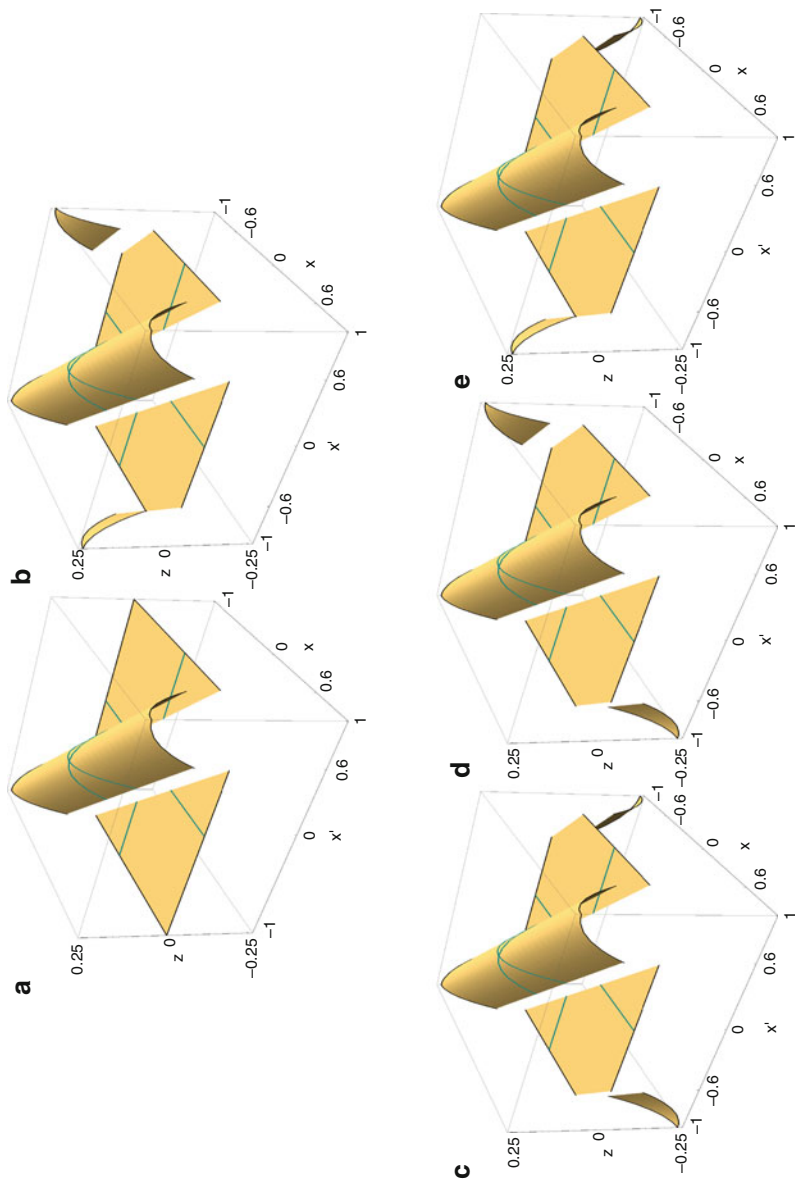


Fig. 2 For the same kernel function $\widehat{C}(x)$ in Fig. 1, we employ bivariate versions of $\widehat{C}(x' - x)$, $\widehat{C}_p(x' - x)$, $\widehat{C}_a(x' - x)$, $\widehat{C}_{pa}(x' - x)$, and $\widehat{C}_{ap}(x' - x)$ with the definition $\widehat{C}_{bc}(x, x') := \widehat{C}_{bc}(x', x)$. Plots of (a) $\widehat{C}(x, x')$, (b) $\widehat{C}_p(x, x')$, (c) $\widehat{C}_a(x, x')$, (d) $\widehat{C}_{pa}(x, x')$, (e) $\widehat{C}_{ap}(x, x')$

agree with $\mathcal{L}u(x, y)$ in the bulk, i.e., for $(x, y) \in (-1 + \delta, 1 - \delta) \times (-1 + \delta, 1 - \delta)$. Furthermore, the operators \mathcal{M}_N and \mathcal{M}_D enforce pure Neumann and Dirichlet BC, respectively:

$$\begin{aligned} \frac{\partial}{\partial n} [(\mathcal{M}_N - c)u](x, \pm 1) &= \frac{\partial}{\partial n} [(\mathcal{M}_N - c)u](\pm 1, y) = 0, \\ (\mathcal{M}_D - c)u(x, \pm 1) &= (\mathcal{M}_D - c)u(\pm 1, y) = 0. \end{aligned}$$

The operators $\mathcal{M}_{ND,ND}$ and $\mathcal{M}_{N,DN}$ enforce mixed Neumann and Dirichlet BC, respectively, in the following way:

$$\begin{aligned} \frac{\partial}{\partial n} [(\mathcal{M}_{ND,ND} - c)u](-1, y) &= \frac{\partial}{\partial n} [(\mathcal{M}_{ND,ND} - c)u](x, -1) = 0, \\ (\mathcal{M}_{ND,ND} - c)u(+1, y) &= (\mathcal{M}_{ND,ND} - c)u(x, +1) = 0, \end{aligned}$$

and

$$\begin{aligned} \frac{\partial}{\partial n} [(\mathcal{M}_{N,DN} - c)u](\pm 1, y) &= \frac{\partial}{\partial n} [(\mathcal{M}_{N,DN} - c)u](x, +1) = 0, \\ (\mathcal{M}_{N,DN} - c)u(x, -1) &= 0. \end{aligned}$$

Proof. The proofs of agreement in the bulk and the verification of BC are given in sections “[The Construction of 2D Operators](#)” and “[Verifying the Boundary Conditions](#),” respectively. □

Remark 1. Although we assume a separable kernel function as in (11), note that we do not impose a separability assumption on the solution $u(x, y)$.

Operators in 1D

The construction in higher dimensions is inspired by the one in 1D. Hence, it is more instructive to provide the construction in 1D. For the convolution present in the governing operators, we use a shorthand notation and define

$$Cu(x) := \int_{\Omega} \widehat{C}(x' - x)u(x')dx'.$$

Furthermore, for each extension type utilized, we define the following operators which will be useful in the exposition. Following the construction in Aksoylu et al. (2017a), we assume that $u, C \in L^2(\Omega)$ and define

$$C_p u(x) := \int_{\Omega} \widehat{C}_p(x' - x)u(x')dx', \quad C_a u(x) := \int_{\Omega} \widehat{C}_a(x' - x)u(x')dx', \tag{12}$$

$$C_{pa} u(x) := \int_{\Omega} \widehat{C}_{pa}(x' - x)u(x')dx', \quad C_{ap} u(x) := \int_{\Omega} \widehat{C}_{ap}(x' - x)u(x')dx'. \tag{13}$$

The only difference in the operators C_p, C_a, C_{pa} , and C_{ap} is the extension utilized for the kernel functions. We prove that the operators agree in the bulk by investigating how the corresponding kernel functions behave in the bulk.

Lemma 2. *Let the kernel function $\widehat{C}(x)$ be in the form*

$$\widehat{C}(x) = \widehat{\chi}_{\delta}(x)\widehat{\mu}(x),$$

where $\widehat{\mu}(x) \in L^2(\widehat{\Omega})$ is even. Let $\widehat{C}_p(x), \widehat{C}_a(x), \widehat{C}_{pa}(x)$, and $\widehat{C}_{ap}(x)$ denote the periodic, antiperiodic, and mixed extensions of $C(x)$ to $\widehat{\Omega}$, respectively. Then,

$$\widehat{C}(x) = \widehat{C}_p(x) = \widehat{C}_a(x) = \widehat{C}_{pa}(x) = \widehat{C}_{ap}(x), \quad x \in (-2 + \delta, 2 - \delta).$$

Furthermore, we have the following agreement in the bulk. Namely, for $x \in (-1 + \delta, 1 - \delta)$,

$$\widehat{C}(x' - x) = \widehat{C}_p(x' - x) = \widehat{C}_a(x' - x) = \widehat{C}_{pa}(x' - x) = \widehat{C}_{ap}(x' - x), \quad x' \in [-1, 1]. \tag{14}$$

Proof. Let us study the definition of $\widehat{C}_p(x)$ given in (10) by explicitly writing the expression of practical kernel (4) as follows:

$$\widehat{C}_p(x) = \begin{cases} \chi_{\delta}(x + 2)\mu(x + 2), & x \in [-2, -1), \\ \chi_{\delta}(x)\mu(x), & x \in [-1, 1], \\ \chi_{\delta}(x - 2)\mu(x - 2), & x \in (1, 2]. \end{cases}$$

Let us closely look at the first expression in the above definition of $\widehat{C}_p(x)$:

$$\widehat{C}_p(x)|_{x \in [-2, -1)} = \chi_{\delta}(x + 2)\mu(x + 2). \tag{15}$$

The expression in (15) is equivalent to

$$\widehat{C}_p(x)|_{x \in [-2, -1)} = \begin{cases} \mu(x + 2), & x + 2 \in (-\delta, \delta) \text{ and } x \in [-2, -1), \\ 0, & x + 2 \notin (-\delta, \delta) \text{ and } x \in [-2, -1). \end{cases} \tag{16}$$

Table 1 The value of each extension of the function C

Interval	$\widehat{C}_p(x)$	$\widehat{C}_a(x)$	$\widehat{C}_{pa}(x)$	$\widehat{C}_{ap}(x)$
$x \in [-2, -2 + \delta)$	$\mu(x + 2)$	$-\mu(x + 2)$	$\mu(x + 2)$	$-\mu(x + 2)$
$x \in [-2 + \delta, -\delta]$	0	0	0	0
$x \in (-\delta, \delta)$	$\mu(x)$	$\mu(x)$	$\mu(x)$	$\mu(x)$
$x \in [\delta, 2 - \delta]$	0	0	0	0
$x \in (2 - \delta, 2]$	$\mu(x - 2)$	$-\mu(x - 2)$	$-\mu(x - 2)$	$\mu(x - 2)$

Due to the following set equivalence

$$\begin{aligned} & \{x : x + 2 \in (-\delta, \delta) \text{ and } x \in [-2, -1)\} \\ &= \{x : x \in (-2 - \delta, -2 + \delta) \cap [-2, -1) = [-2, -2 + \delta)\}, \end{aligned}$$

the expression (16) reduces to

$$\widehat{C}_p(x)|_{x \in [-2, -1)} = \begin{cases} \mu(x + 2), & x \in [-2, -2 + \delta), \\ 0, & x \in [-2 + \delta, -1). \end{cases} \tag{17}$$

Similar to (17), for $x \in (1, 2]$, we have

$$\widehat{C}_p(x)|_{x \in (1, 2]} = \begin{cases} 0, & x \in (1, 2 - \delta], \\ \mu(x - 2), & x \in (2 - \delta, 2]. \end{cases} \tag{18}$$

Combining (17) and (18), for $x \in [-2, 2]$, we obtain the expression for $\widehat{C}_p(x)$. Similarly, we obtain the expressions for the antiperiodic and the mixed extensions. We collect all the expressions in Table 1.

Clearly, all of the extensions agree for $x \in [-2 + \delta, 2 - \delta]$. Also, see Figs. 1 and 2. □

Using (14), we immediately obtain the following equivalence of operators in the bulk:

$$Cu(x) = C_p u(x) = C_a u(x) = C_{pa} u(x) = C_{ap} u(x), \quad x \in (-1 + \delta, 1 - \delta). \tag{19}$$

Even and odd parts of a univariate function $u(x)$ are used in the governing operators $\mathcal{M}_N, \mathcal{M}_D, \mathcal{M}_{ND},$ and \mathcal{M}_{DN} . We define the orthogonal projections that give the even and odd parts, respectively, of a univariate function by $P_e, P_o : L^2(\Omega) \rightarrow L^2(\Omega)$, whose definitions are

$$P_e u(x) := \frac{u(x) + u(-x)}{2}, \quad P_o u(x) := \frac{u(x) - u(-x)}{2}.$$

We present a commutativity property that allows us to identify the kernel functions associated with the operators \mathcal{M}_N and \mathcal{M}_D .

Lemma 3.

$$\mathcal{C}_p P_e = P_e \mathcal{C}_p, \quad \mathcal{C}_p P_o = P_o \mathcal{C}_p, \quad \mathcal{C}_a P_e = P_e \mathcal{C}_a, \quad \mathcal{C}_a P_o = P_o \mathcal{C}_a. \quad (20)$$

Proof. We present the proof for $\mathcal{C}_p P_e = P_e \mathcal{C}_p$. The other results easily follow. We recall the definition of $\mathcal{C}_p u(x)$ in (12). We explicitly write $P_e \mathcal{C}_p u(x)$, and the result follows by utilizing the evenness of $\widehat{\mathcal{C}}_p$ and a change of variable:

$$\begin{aligned} P_e \mathcal{C}_p u(x) &= \frac{1}{2} \left(\int_{\Omega} \widehat{\mathcal{C}}_p(x' - x) u(x') dx' + \int_{\Omega} \widehat{\mathcal{C}}_p(x' + x) u(x') dx' \right) \\ &= \frac{1}{2} \left(\int_{\Omega} \widehat{\mathcal{C}}_p(x' - x) u(x') dx' + \int_{\Omega} \widehat{\mathcal{C}}_p(x' - x) u(-x') dx' \right) \\ &= \int_{\Omega} \widehat{\mathcal{C}}_p(x' - x) P_e u(x') dx' \\ &= \mathcal{C}_p P_e u(x). \end{aligned}$$

□

Remark 2. The above commutativity property plays an important role in determining the spectrum of the operators \mathcal{M}_N and \mathcal{M}_D ; see Aksoylu and Kaya (2018). It also helps in identifying the associated kernel functions; see (22) and (23). Note that the above commutativity property does not hold for the operators \mathcal{C}_{pa} and \mathcal{C}_{ap} . Identification of the associated kernel functions can be done by direct manipulation.

Theorem 2 (Main Theorem in 1D). Let $c = \int_{\Omega} C(x') dx'$. The following operators \mathcal{M}_N , \mathcal{M}_D , \mathcal{M}_{ND} , and \mathcal{M}_{DN} defined by

$$\begin{aligned} (\mathcal{M}_D - c)u(x) &:= - \int_{\Omega} [\widehat{\mathcal{C}}_a(x' - x) P_e u(x') + \widehat{\mathcal{C}}_p(x' - x) P_o u(x')] dx', \\ (\mathcal{M}_N - c)u(x) &:= - \int_{\Omega} [\widehat{\mathcal{C}}_p(x' - x) P_e u(x') + \widehat{\mathcal{C}}_a(x' - x) P_o u(x')] dx', \\ (\mathcal{M}_{ND} - c)u(x) &:= - \int_{\Omega} [\widehat{\mathcal{C}}_{ap}(x' - x) P_e u(x') + \widehat{\mathcal{C}}_{pa}(x' - x) P_o u(x')] dx', \\ (\mathcal{M}_{DN} - c)u(x) &:= - \int_{\Omega} [\widehat{\mathcal{C}}_{pa}(x' - x) P_e u(x') + \widehat{\mathcal{C}}_{ap}(x' - x) P_o u(x')] dx' \end{aligned}$$

agree with $\mathcal{L}u(x)$ in the bulk, i.e., for $x \in (-1+\delta, 1-\delta)$. Furthermore, the operators \mathcal{M}_N and \mathcal{M}_D enforce pure Neumann and Dirichlet BC, respectively:

$$\begin{aligned} \frac{d}{dx}[(\mathcal{M}_N - c)u](\pm 1) &= 0, \\ (\mathcal{M}_D - c)u(\pm 1) &= 0. \end{aligned}$$

The operators \mathcal{M}_{ND} and \mathcal{M}_{DN} enforce mixed Neumann and Dirichlet BC, respectively:

$$\begin{aligned} (\mathcal{M}_{ND} - c)u(+1) &= \frac{d}{dx}[(\mathcal{M}_{ND} - c)u](-1) = 0, \\ (\mathcal{M}_{DN} - c)u(-1) &= \frac{d}{dx}[(\mathcal{M}_{DN} - c)u](+1) = 0. \end{aligned}$$

We define the operators \mathcal{M}_N , \mathcal{M}_D , \mathcal{M}_{ND} , and \mathcal{M}_{DN} as bounded, linear operators. More precisely, \mathcal{M}_D , \mathcal{M}_N , \mathcal{M}_{ND} , $\mathcal{M}_{DN} \in L(X, X)$ where $X = L^2(\Omega) \cap C^1(\partial\Omega)$. For \mathcal{M}_D , the choice of X can be relaxed as $L^2(\Omega) \cap C^0(\partial\Omega)$. This choice is implied when we study \mathcal{M}_D . The assumptions for the operators \mathcal{M}_{ND} and \mathcal{M}_{DN} are also implied in a similar way.

Imposing Neumann (also periodic and antiperiodic) BC requires differentiation. For technical details regarding differentiation under the integral sign, see the discussion on the Leibniz rule in Aksoylu et al. (2017) whose proof relies on the Lebesgue dominated convergence theorem. In addition, the limit in the definition of the Dirichlet BC can be interchanged with the integral sign, again by the Lebesgue dominated convergence theorem.

Remark 3. When we assume homogeneous Neumann and Dirichlet BC on u , then the operators \mathcal{M}_N and \mathcal{M}_D enforce homogeneous Dirichlet and Neumann BC, respectively. More precisely, for $u(\pm 1) = 0$ and $u'(\pm 1) = 0$, we obtain $\frac{d}{dx}\mathcal{M}_N u(\pm 1) = 0$ and $\mathcal{M}_D u(\pm 1) = 0$, respectively. The same line of argument applies to the operators \mathcal{M}_{ND} and \mathcal{M}_{DN} .

Using the operators \mathcal{C}_p , \mathcal{C}_a , \mathcal{C}_{pa} , and \mathcal{C}_{ap} given in (12) and (13), we can express the operators \mathcal{M}_N , \mathcal{M}_D , \mathcal{M}_{ND} , and \mathcal{M}_{DN} in the following way:

$$\begin{aligned} (\mathcal{M}_N - c) &= -(\mathcal{C}_p P_e + \mathcal{C}_a P_o), & (\mathcal{M}_D - c) &= -(\mathcal{C}_a P_e + \mathcal{C}_p P_o), \\ (\mathcal{M}_{ND} - c) &= -(\mathcal{C}_{ap} P_e + \mathcal{C}_{pa} P_o), & (\mathcal{M}_{DN} - c) &= -(\mathcal{C}_{pa} P_e + \mathcal{C}_{ap} P_o). \end{aligned}$$

Using the commutativity property (20), we arrive at the following representation:

$$(\mathcal{M}_N - c) = -(P_e \mathcal{C}_p + P_o \mathcal{C}_a), \quad (\mathcal{M}_D - c) = -(P_e \mathcal{C}_a + P_o \mathcal{C}_p).$$

Now, we can identify the kernel functions associated with operators \mathcal{M}_N and \mathcal{M}_D :

$$\begin{aligned}(\mathcal{M}_N - c)u(x) &= -\int_{\Omega} K_N(x, x')u(x')dx', \\(\mathcal{M}_D - c)u(x) &= -\int_{\Omega} K_D(x, x')u(x')dx',\end{aligned}\tag{21}$$

where

$$K_N(x, x') := \frac{1}{2} \{ [\widehat{C}_p(x' - x) + \widehat{C}_p(x' + x)] + [\widehat{C}_a(x' - x) - \widehat{C}_a(x' + x)] \},\tag{22}$$

$$K_D(x, x') := \frac{1}{2} \{ [\widehat{C}_a(x' - x) + \widehat{C}_a(x' + x)] + [\widehat{C}_p(x' - x) - \widehat{C}_p(x' + x)] \}.\tag{23}$$

We also want to identify the integrands associated with the operators \mathcal{M}_{ND} and \mathcal{M}_{DN} . We proceed by direct manipulation. By writing P_e and P_o explicitly and utilizing a simple change of variable, we arrive at the following expressions:

$$\begin{aligned}(\mathcal{M}_{ND} - c)u(x) &= -\int_{\Omega} K_{ND}(x, x')u(x')dx', \\(\mathcal{M}_{DN} - c)u(x) &= -\int_{\Omega} K_{DN}(x, x')u(x')dx',\end{aligned}$$

where

$$\begin{aligned}K_{ND}(x, x') &:= \frac{1}{2} \{ [\widehat{C}_{pa}(x' - x) + \widehat{C}_{pa}(x' + x)] + [\widehat{C}_{ap}(x' - x) - \widehat{C}_{ap}(x' + x)] \}, \\K_{DN}(x, x') &:= \frac{1}{2} \{ [\widehat{C}_{ap}(x' - x) + \widehat{C}_{ap}(x' + x)] + [\widehat{C}_{pa}(x' - x) - \widehat{C}_{pa}(x' + x)] \}.\end{aligned}$$

In order to align with the construction given in Aksoylu et al. (2017a), we assume that

$$\widehat{C}(x), \widehat{C}_a(x), \widehat{C}_p(x), \widehat{C}_{pa}(x), \text{ and } \widehat{C}_{ap}(x) \in L^2(\widehat{\Omega}).\tag{24}$$

Remark 4. The boundedness of \mathcal{M}_N , \mathcal{M}_D , \mathcal{M}_{ND} , and \mathcal{M}_{DN} follow from the choices of (5) and (24). In addition, all of them fall into the class of integral operators; hence, their self-adjointness follows from the fact that the corresponding kernels are symmetric (due to evenness of C), i.e., $K_{BC}(x, x') = K_{BC}(x', x)$ and $BC \in \{N, D, ND, DN\}$. The cases of $BC \in \{ND, DN\}$ are more involved than the rest. One useful identity is $\widehat{C}_{ap}(x' - x) = \widehat{C}_{pa}(-x' + x)$.

In the upcoming proofs, we want to report a minor caveat. We use $\widehat{C}_a(x' + 1) = -\widehat{C}_a(x' - 1)$ which holds for $x' \neq 0$. For $x' = 0$, i.e., $\widehat{C}_a(x' + 1) = C(1) \neq -C(-1) = -\widehat{C}_a(x' - 1)$. Since $x' = 0$ is only a point, it does not change the value of the integral. We choose not to point it out each time we run into this case.

Proof (Proof of Theorem 2). The key observation that leads to the agreement of the operators $\mathcal{M}_N, \mathcal{M}_D, \mathcal{M}_{ND}$, and \mathcal{M}_{DN} with the operator \mathcal{L} is the agreement of kernel functions in (14). The property (14) leads to the equivalence (19). Hence, we arrive at the following equivalence for $x \in (-1 + \delta, 1 - \delta)$:

$$\begin{aligned}
 (\mathcal{L} - c) &= -\mathcal{C} \\
 &= -\mathcal{C}(P_e + P_o) \quad (\text{using } u = P_e u + P_o u) \\
 &= -(\mathcal{C}P_e + \mathcal{C}P_o) \\
 &= -(\mathcal{C}_p P_e + \mathcal{C}_a P_o) \quad (\text{using (19)}) \\
 &=: (\mathcal{M}_N - c).
 \end{aligned} \tag{25}$$

Similar to (25), we can show that the other operators agree in the bulk as well:

$$\begin{aligned}
 (\mathcal{L} - c) &= -\mathcal{C} \\
 &= -(\mathcal{C}_a P_e + \mathcal{C}_p P_o) \quad =: (\mathcal{M}_D - c)
 \end{aligned} \tag{26}$$

$$= -(\mathcal{C}_{ap} P_e + \mathcal{C}_{pa} P_o) \quad =: (\mathcal{M}_{ND} - c) \tag{27}$$

$$= -(\mathcal{C}_{pa} P_e + \mathcal{C}_{ap} P_o) \quad =: (\mathcal{M}_{DN} - c) \tag{28}$$

First, we prove that the operators \mathcal{M}_N and \mathcal{M}_D enforce pure Neumann and Dirichlet BC, respectively. Next, we will prove that the operators \mathcal{M}_{ND} and \mathcal{M}_{DN} enforce mixed Neumann and Dirichlet BC, respectively.

- **The operator \mathcal{M}_N :** First we remove the points at which the partial derivative of $K_N(x, x')$ does not exist from the set of integration. Note that such points form a set of measure zero and, hence, do not affect the value of the integral. We differentiate both sides of (21). In Aksoylyu et al. (2017), we had proved that the differentiation in the definition of the Neumann BC can interchange with the integral. We can differentiate the integrand $K_N(x, x')$ piecewise and obtain

$$\frac{d}{dx} [(\mathcal{M}_N - c)u](x) = - \int_{\Omega} \frac{\partial K_N}{\partial x}(x, x') u(x') dx', \tag{29}$$

where

$$\frac{\partial K_N}{\partial x}(x, x') = \frac{1}{2} \{ [-\widehat{C}'_p(x' - x) + \widehat{C}'_p(x' + x)] + [-\widehat{C}'_a(x' - x) - \widehat{C}'_a(x' + x)] \}.$$

We check the boundary values by plugging $x = \pm 1$ in (29).

$$\frac{d}{dx}[(\mathcal{M}_N - c)u](\pm 1) = - \int_{\Omega} \frac{\partial K_N}{\partial x}(\pm 1, x')u(x')dx'. \quad (30)$$

The functions \widehat{C}'_p and \widehat{C}'_a are 2-periodic and 2-antiperiodic because they are the derivatives of 2-periodic and 2-antiperiodic functions, respectively. Hence,

$$\widehat{C}'_p(x' \mp 1) = \widehat{C}'_p(x' \pm 1) \quad \text{and} \quad \widehat{C}'_a(x' \mp 1) = -\widehat{C}'_a(x' \pm 1).$$

Hence, the integrand in (30) vanishes, i.e.,

$$\frac{\partial K_N}{\partial x}(\pm 1, x') = 0.$$

Therefore, we arrive at

$$\frac{d}{dx}\mathcal{M}_N u(\pm 1) = cu'(\pm 1).$$

When we assume that u satisfies homogeneous Neumann BC, i.e., $u'(\pm 1) = 0$, we conclude that the operator \mathcal{M}_N enforces homogeneous Neumann BC as well.

- **The operator \mathcal{M}_D :** By the Lebesgue dominated convergence theorem, the limit in the definition of the Dirichlet BC can be interchanged with the integral. Now, we check the boundary values by plugging $x = \pm 1$ in (23).

$$(\mathcal{M}_D - c)u(\pm 1) = - \int_{\Omega} K_D(\pm 1, x')u(x')dx'. \quad (31)$$

Since \widehat{C}_p and \widehat{C}_a are 2-periodic and 2-antiperiodic, respectively, we have

$$\widehat{C}_p(x' \mp 1) = \widehat{C}_p(x' \pm 1) \quad \text{and} \quad \widehat{C}_a(x' \mp 1) = -\widehat{C}_a(x' \pm 1).$$

Hence, the integrand in (31) vanishes, i.e., $K_D(\pm 1, x') = 0$. Therefore, we arrive at

$$\mathcal{M}_D u(\pm 1) = cu(\pm 1).$$

When we assume that u satisfies homogeneous Dirichlet BC, i.e., $u(\pm 1) = 0$, we conclude that the operator \mathcal{M}_D enforces homogeneous Dirichlet BC as well.

- **The operator \mathcal{M}_{ND} :** First we prove that $\mathcal{C}_{ap}P_e u(+1) = 0$. We use a change of variable in the second piece.

$$\mathcal{C}_{ap}P_e u(+1) = \frac{1}{2} \left(\int_{\Omega} \widehat{C}_{ap}(x' - 1)P_e u(x')dx' + \int_{\Omega} \widehat{C}_{ap}(-x' - 1)P_e u(x')dx' \right).$$

Then, we split the integrals into two parts as follows:

$$\begin{aligned} C_{\text{ap}} P_e u(+1) &= \frac{1}{2} \int_{-1}^0 [\widehat{C}_{\text{ap}}(x' - 1) + \widehat{C}_{\text{ap}}(-x' - 1)] P_e u(x') dx' \\ &\quad + \frac{1}{2} \int_0^1 [\widehat{C}_{\text{ap}}(x' - 1) + \widehat{C}_{\text{ap}}(-x' - 1)] P_e u(x') dx'. \end{aligned} \quad (32)$$

For $x' \in [-1, 0]$, we have $x' - 1 \in [-2, -1]$. By using the definition of \widehat{C}_{ap} and the evenness of C , we obtain

$$\widehat{C}_{\text{ap}}(x' - 1) = -\widehat{C}_{\text{ap}}(x' + 1) = -C(x' + 1) = -C(-x' - 1) = -\widehat{C}_{\text{ap}}(-x' - 1). \quad (33)$$

For $x' \in [0, 1]$, we have $x' - 1 \in [-1, 0]$. By using the definition of \widehat{C}_{ap} and the evenness of C , we obtain

$$\widehat{C}_{\text{ap}}(x' - 1) = C(x' - 1) = C(-x' + 1) = -\widehat{C}_{\text{ap}}(-x' - 1). \quad (34)$$

Combining (33) and (34) with (32), we conclude that $C_{\text{ap}} P_e u(+1) = 0$. Similarly, we can conclude that $C_{\text{pa}} P_o u(+1) = 0$. Consequently, we arrive at

$$C_{\text{ND}} u(+1) = 0.$$

We prove that $\frac{d}{dx} C_{\text{pa}} P_o u(-1) = 0$. We use a change of variable in the second piece.

$$\begin{aligned} \frac{d}{dx} C_{\text{pa}} P_o u(-1) &= -\frac{1}{2} \left(\int_{\Omega} \widehat{C}'_{\text{pa}}(x' + 1) P_o u(x') dx' \right. \\ &\quad \left. - \int_{\Omega} \widehat{C}'_{\text{pa}}(-x' + 1) P_o u(x') dx' \right). \end{aligned}$$

Then, we split the integrals into two parts as follows:

$$\begin{aligned} \frac{d}{dx} C_{\text{pa}} P_o u(-1) &= -\frac{1}{2} \int_{-1}^0 [\widehat{C}'_{\text{pa}}(x' + 1) - \widehat{C}'_{\text{pa}}(-x' + 1)] P_o u(x') dx' \\ &\quad - \frac{1}{2} \int_0^1 [\widehat{C}'_{\text{pa}}(x' + 1) - \widehat{C}'_{\text{pa}}(-x' + 1)] P_o u(x') dx'. \end{aligned} \quad (35)$$

For $x' \in [-1, 0]$, we have $x' + 1 \in [0, 1]$. By using the definition of \widehat{C}_{pa} and the oddness of C' , we obtain

$$\widehat{C}'_{\text{pa}}(x' + 1) = C'(x' + 1) = -C'(-x' - 1) = -\widehat{C}'_{\text{pa}}(-x' - 1) = \widehat{C}'_{\text{pa}}(-x' + 1). \quad (36)$$

For $x' \in [0, 1]$, we have $x' + 1 \in [1, 2]$. By using the definition of \widehat{C}_{pa} and the oddness of C' , we obtain

$$\widehat{C}'_{\text{pa}}(x' + 1) = -\widehat{C}'_{\text{pa}}(x' - 1) = -C'(x' - 1) = C'(-x' + 1) = \widehat{C}'_{\text{pa}}(-x' + 1). \quad (37)$$

Combining (36) and (37) with (35), we conclude that $\frac{d}{dx}C_{\text{pa}}P_o u(-1) = 0$. Similarly, we can conclude that $\frac{d}{dx}C_{\text{ap}}P_e u(-1) = 0$. Consequently, we arrive at

$$\frac{d}{dx}C_{\text{ND}}u(-1) = 0.$$

- **The operator \mathcal{M}_{DN} :** The proof is similar to the case of \mathcal{M}_{ND} . □

Remark 5. As we prepare to construct the operators in 2D, it is useful to explicitly denote the variable x' on which P_e and P_o act in the following way.

$$C_{\text{N}}u(x) := (C_{\text{p}}P_{e,x'} + C_{\text{a}}P_{o,x'})u(x)$$

$$C_{\text{D}}u(x) := (C_{\text{a}}P_{e,x'} + C_{\text{p}}P_{o,x'})u(x)$$

$$C_{\text{ND}}u(x) := (C_{\text{ap}}P_{e,x'} + C_{\text{pa}}P_{o,x'})u(x)$$

$$C_{\text{DN}}u(x) := (C_{\text{pa}}P_{e,x'} + C_{\text{ap}}P_{o,x'})u(x).$$

Consequently, checking if the operators enforce the BC reduces to obtaining

$$\begin{aligned} \frac{d}{dx}C_{\text{N}}u(\pm 1) = 0, \quad C_{\text{D}}u(\pm 1) &= 0 \\ \frac{d}{dx}C_{\text{ND}}u(-1) = 0, \quad C_{\text{ND}}u(+1) &= 0 \\ \frac{d}{dx}C_{\text{DN}}u(+1) = 0, \quad C_{\text{DN}}u(-1) &= 0. \end{aligned} \quad (38)$$

The Construction of 2D Operators

For the convolution present in the governing operators, we use a shorthand notation and define the operator

$$Cu(x, y) := \iint_{\Omega} \widehat{C}(x' - x, y' - y)u(x', y')dx'dy'. \quad (39)$$

We also define the following auxiliary operators that act on a bivariate function.

$$\mathcal{X}_E u(x, y) := \int_{\Omega} \widehat{X}_E(x' - x)u(x', y)dx', \quad \mathcal{Y}_E u(x, y) := \int_{\Omega} \widehat{Y}_E(y' - y)u(x, y')dy',$$

where $E \in \{p, a, pa, ap\}$.

Using the separability assumption (11) on the kernel function, we have the following:

$$\widehat{C}(x, y) = \widehat{X}(x)\widehat{Y}(y). \tag{40}$$

The separability of the kernel function leads to the following important property. Using (40) and the Fubini theorem, we rewrite the operator \mathcal{C} in (39).

$$\begin{aligned} \mathcal{C}u(x, y) &= \iint_{\Omega} \widehat{X}(x' - x)\widehat{Y}(y' - y)u(x', y')dx'dy' \\ &= \int_{\Omega} \widehat{X}(x' - x)\left(\int_{\Omega} \widehat{Y}(y' - y)u(x', y')dy'\right)dx' \\ &= \int_{\Omega} \widehat{X}(x' - x)\left(\mathcal{Y}u(x', y)\right)dx' \\ &= \mathcal{X}(\mathcal{Y}u)(x, y) \end{aligned} \tag{41}$$

In other words, we proved that \mathcal{C} can be decomposed into a product of two 1D operators where the action of \mathcal{X} and \mathcal{Y} is on the variables x and y , respectively. Furthermore, a change in the order of integration leads to

$$\mathcal{C}u(x, y) = \mathcal{Y}\mathcal{X}u(x, y). \tag{42}$$

Similar to (19), we also obtain the following equivalence of operators in the bulk. For fixed y_0 , we have

$$\begin{aligned} \mathcal{X}u(x, y_0) &= \mathcal{X}_p u(x, y_0) = \mathcal{X}_a u(x, y_0) \\ &= \mathcal{X}_{pa} u(x, y_0) = \mathcal{X}_{ap} u(x, y_0), \quad x \in (-1 + \delta, 1 - \delta). \end{aligned}$$

Also, for fixed x_0 , we have

$$\begin{aligned} \mathcal{Y}u(x_0, y) &= \mathcal{Y}_p u(x_0, y) = \mathcal{Y}_a u(x_0, y) = \mathcal{Y}_{pa} u(x_0, y) \\ &= \mathcal{Y}_{ap} u(x_0, y) \quad y \in (-1 + \delta, 1 - \delta). \end{aligned}$$

The choice made in (25) leads to the construction of the operator that enforces pure Neumann BC in the x - and y -variable as follows:

$$\mathcal{X}_N := \mathcal{X}_p P_{e,x'} + \mathcal{X}_a P_{o,x'} \quad (\text{in the } x\text{-variable}) \quad (43)$$

$$\mathcal{Y}_N := \mathcal{Y}_p P_{e,y'} + \mathcal{Y}_a P_{o,y'} \quad (\text{in the } y\text{-variable}). \quad (44)$$

Similarly, the choice made in (26) leads to the construction of the operator that enforces pure Dirichlet BC in the x - and y -variable as follows:

$$\mathcal{X}_D := \mathcal{X}_a P_{e,x'} + \mathcal{X}_p P_{o,x'} \quad (\text{in the } x\text{-variable}) \quad (45)$$

$$\mathcal{Y}_D := \mathcal{Y}_a P_{e,y'} + \mathcal{Y}_p P_{o,y'} \quad (\text{in the } y\text{-variable}). \quad (46)$$

Similarly, the choices made in (27) and (28) lead to the construction of the operators that enforce mixed Neumann-Dirichlet and Dirichlet-Neumann BC in the x - and y -variable as follows:

$$\mathcal{X}_{ND} := \mathcal{X}_{ap} P_{e,x'} + \mathcal{X}_{pa} P_{o,x'} \quad (\text{in the } x\text{-variable}) \quad (47)$$

$$\mathcal{Y}_{ND} := \mathcal{Y}_{ap} P_{e,y'} + \mathcal{Y}_{pa} P_{o,y'} \quad (\text{in the } y\text{-variable}) \quad (48)$$

$$\mathcal{X}_{DN} := \mathcal{X}_{pa} P_{e,x'} + \mathcal{X}_{ap} P_{o,x'} \quad (\text{in the } x\text{-variable}) \quad (49)$$

$$\mathcal{Y}_{DN} := \mathcal{Y}_{pa} P_{e,y'} + \mathcal{Y}_{ap} P_{o,y'} \quad (\text{in the } y\text{-variable}). \quad (50)$$

We want to construct an operator that enforces pure Neumann BC on the square. We make the choice that gives the 1D Neumann operator both in x - and y -variables. Hence, combining (43) and (44), we define the 2D pure Neumann operator as

$$(\mathcal{M}_N - c) := -\mathcal{X}_N \mathcal{Y}_N = -(\mathcal{X}_p P_{e,x'} + \mathcal{X}_a P_{o,x'}) (\mathcal{Y}_p P_{e,y'} + \mathcal{Y}_a P_{o,y'}). \quad (51)$$

Similarly, combining (45) and (46), we define the 2D pure Dirichlet operator as

$$(\mathcal{M}_D - c) := -\mathcal{X}_D \mathcal{Y}_D = -(\mathcal{X}_a P_{e,x'} + \mathcal{X}_p P_{o,x'}) (\mathcal{Y}_a P_{e,y'} + \mathcal{Y}_p P_{o,y'}). \quad (52)$$

Similarly, combining (47), (48), (49), and (50), we define the 2D mixed operators as follows:

$$(\mathcal{M}_{ND,ND} - c) := -\mathcal{X}_{ND} \mathcal{Y}_{ND} = -(\mathcal{X}_{ap} P_{e,x'} + \mathcal{X}_{pa} P_{o,x'}) (\mathcal{Y}_{ap} P_{e,y'} + \mathcal{Y}_{pa} P_{o,y'}) \quad (53)$$

$$(\mathcal{M}_{N,DN} - c) := -\mathcal{X}_N \mathcal{Y}_{DN} = -(\mathcal{X}_p P_{e,x'} + \mathcal{X}_a P_{o,x'}) (\mathcal{Y}_{pa} P_{e,y'} + \mathcal{Y}_{ap} P_{o,y'}). \quad (54)$$

Recalling (2), we immediately see that the operator \mathcal{L} agrees in the bulk with the given operators above. Namely,

$$\begin{aligned}
 (\mathcal{L} - c) &= -\mathcal{C} = (\mathcal{M}_N - c) \\
 (\mathcal{L} - c) &= -\mathcal{C} = (\mathcal{M}_D - c) \\
 (\mathcal{L} - c) &= -\mathcal{C} = (\mathcal{M}_{ND,ND} - c) \\
 (\mathcal{L} - c) &= -\mathcal{C} = (\mathcal{M}_{N,DN} - c).
 \end{aligned}$$

Remark 6. The operator \mathcal{C} in (2) utilizes a 2D computational domain which is indicated by the integration variable $dx'dy' = d(x', y')$. We can show the construction of each operator by paying attention to the computational domain of each operator and rearranging (41) using the agreement of operators in the bulk in the following way:

$$\begin{aligned}
 \mathcal{C} &= \mathcal{C}I_{x',y'} = \mathcal{X}I_{x'} \mathcal{Y}I_{y'} \\
 &= \mathcal{X}(P_{e,x'} + P_{o,x'}) \mathcal{Y}(P_{e,y'} + P_{o,y'}) \\
 &= (\mathcal{X}P_{e,x'} + \mathcal{X}P_{o,x'}) (\mathcal{Y}P_{e,y'} + \mathcal{Y}P_{o,y'}) \\
 &= (\mathcal{X}_p P_{e,x'} + \mathcal{X}_a P_{o,x'}) (\mathcal{Y}_p P_{e,y'} + \mathcal{Y}_a P_{o,y'}) \quad =: -(\mathcal{M}_N - c) \\
 &= (\mathcal{X}_a P_{e,x'} + \mathcal{X}_p P_{o,x'}) (\mathcal{Y}_a P_{e,y'} + \mathcal{Y}_p P_{o,y'}) \quad =: -(\mathcal{M}_D - c) \\
 &= (\mathcal{X}_{ap} P_{e,x'} + \mathcal{X}_{ap} P_{o,x'}) (\mathcal{Y}_{ap} P_{e,y'} + \mathcal{Y}_{ap} P_{o,y'}) \quad =: -(\mathcal{M}_{ND,ND} - c) \quad (55) \\
 &= (\mathcal{X}_p P_{e,x'} + \mathcal{X}_a P_{o,x'}) (\mathcal{Y}_{ap} P_{e,y'} + \mathcal{Y}_{pa} P_{o,y'}) \quad =: -(\mathcal{M}_{N,DN} - c). \quad (56)
 \end{aligned}$$

We construct the operators in higher dimensions by using the corresponding rearrangement; see section “Operators in Higher Dimensions” for the 3D construction. In addition, the 2D decomposition operator $I_{x',y'}$ given in (9) is indeed the product of the 1D decomposition operators $I_{x'}$ and $I_{y'}$ given in (8). More precisely,

$$\begin{aligned}
 I_{x',y'} &= I_{x'} I_{y'} \\
 &= (P_{e,x'} + P_{o,x'})(P_{e,y'} + P_{o,y'}) \\
 &= P_{e,x'}P_{e,y'} + P_{e,x'}P_{o,y'} + P_{o,x'}P_{e,y'} + P_{o,x'}P_{o,y'}.
 \end{aligned}$$

Verifying the Boundary Conditions

The operators $(\mathcal{M}_N - c)$, $(\mathcal{M}_D - c)$, $(\mathcal{M}_{ND,ND} - c)$, and $(\mathcal{M}_{N,DN} - c)$ given in (51), (52), (53), and (54), respectively, are the product of two 1D operators. As we mentioned, the limit in the definition of the BC can be interchanged with the integral sign due to the Lebesgue dominated convergence theorem and the Leibniz rule. Then, using the change in the order of integration as in (42) and (38), we can prove that the pure and mixed Neumann and Dirichlet BC are enforced.

First, we prove that the operators \mathcal{M}_N and \mathcal{M}_D enforce pure Neumann and pure Dirichlet BC in 2D:

$$\begin{aligned}\frac{\partial}{\partial n} [(\mathcal{M}_N - c) u](x, +1) &= - \left(\frac{\partial}{\partial y} \mathcal{X}_N \mathcal{Y}_N \right) u(x, +1) \\ &= - \mathcal{X}_N \left(\frac{d}{dy} \mathcal{Y}_N \right) u(x, +1) = 0\end{aligned}$$

$$\begin{aligned}\frac{\partial}{\partial n} [(\mathcal{M}_N - c) u](x, -1) &= \left(\frac{\partial}{\partial y} \mathcal{X}_N \mathcal{Y}_N \right) u(x, -1) \\ &= \mathcal{X}_N \left(\frac{d}{dy} \mathcal{Y}_N \right) u(x, -1) = 0\end{aligned}$$

$$\begin{aligned}\frac{\partial}{\partial n} [(\mathcal{M}_N - c) u](+1, y) &= - \left(\frac{\partial}{\partial x} \mathcal{Y}_N \mathcal{X}_N \right) u(+1, y) \\ &= - \mathcal{Y}_N \left(\frac{d}{dx} \mathcal{X}_N \right) u(+1, y) = 0\end{aligned}$$

$$\begin{aligned}\frac{\partial}{\partial n} [(\mathcal{M}_N - c) u](-1, y) &= \left(\frac{\partial}{\partial x} \mathcal{Y}_N \mathcal{X}_N \right) u(-1, y) \\ &= \mathcal{Y}_N \left(\frac{d}{dx} \mathcal{X}_N \right) u(-1, y) = 0.\end{aligned}$$

$$(\mathcal{M}_D - c)u(x, \pm 1) = -\mathcal{X}_D \mathcal{Y}_D u(x, \pm 1) = 0$$

$$(\mathcal{M}_D - c)u(\pm 1, y) = -\mathcal{Y}_D \mathcal{X}_D u(\pm 1, y) = 0.$$

Then, we prove that the operator $\mathcal{M}_{ND, ND}$ enforces mixed (2+2) Neumann-Dirichlet, i.e., the West and South edges have Neumann and the East and North edges have Dirichlet BC:

$$\begin{aligned}\frac{\partial}{\partial n} [(\mathcal{M}_{ND, ND} - c) u](-1, y) &= \left(\frac{\partial}{\partial x} \mathcal{Y}_{ND} \mathcal{X}_{ND} \right) u(-1, y) \\ &= \mathcal{Y}_{ND} \left(\frac{d}{dx} \mathcal{X}_{ND} \right) u(-1, y) = 0\end{aligned}$$

$$\begin{aligned}\frac{\partial}{\partial n} [(\mathcal{M}_{ND, ND} - c) u](x, -1) &= \left(\frac{\partial}{\partial y} \mathcal{X}_{ND} \mathcal{Y}_{ND} \right) u(x, -1) \\ &= \mathcal{X}_{ND} \left(\frac{d}{dy} \mathcal{Y}_{ND} \right) u(x, -1) = 0.\end{aligned}$$

$$(\mathcal{M}_{ND, ND} - c)u(+1, y) = -\mathcal{X}_{ND} \mathcal{Y}_{ND} u(+1, y) = 0$$

$$(\mathcal{M}_{ND, ND} - c)u(x, +1) = -\mathcal{X}_{ND} \mathcal{Y}_{ND} u(x, +1) = 0.$$

Finally, we prove that the operator $\mathcal{M}_{N, DN}$ enforces mixed Neumann-Dirichlet (3+1), i.e., the East, West, and North edges have Neumann and the South edge have Dirichlet BC:

$$\begin{aligned} \frac{\partial}{\partial n} [(\mathcal{M}_{N, DN} - c)u](+1, y) &= -\left(\frac{\partial}{\partial x} \mathcal{Y}_{DN} \mathcal{X}_N\right) u(+1, y) \\ &= \mathcal{Y}_{DN} \left(\frac{d}{dx} \mathcal{X}_N\right) u(+1, y) = 0 \\ \frac{\partial}{\partial n} [(\mathcal{M}_{N, DN} - c)u](-1, y) &= \left(\frac{\partial}{\partial x} \mathcal{Y}_{DN} \mathcal{X}_N\right) u(-1, y) \\ &= \mathcal{Y}_{DN} \left(\frac{d}{dx} \mathcal{X}_N\right) u(-1, y) = 0 \\ \frac{\partial}{\partial n} [(\mathcal{M}_{N, DN} - c)u](x, +1) &= -\left(\frac{\partial}{\partial y} \mathcal{X}_N \mathcal{Y}_{DN}\right) u(x, +1) \\ &= -\mathcal{X}_N \left(\frac{d}{dy} \mathcal{Y}_{DN}\right) u(x, +1) = 0. \\ (\mathcal{M}_{N, DN} - c)u(x, -1) &= -\mathcal{X}_N \mathcal{Y}_{DN} u(x, -1) = 0. \end{aligned}$$

Operators in Higher Dimensions

Let us consider the convolution in 3D and the domain be $\Omega := [-1, 1] \times [-1, 1] \times [-1, 1]$. We define the convolution in 3D similarly using notation in (39).

$$Cu(x, y) = \iiint_{\Omega} \widehat{C}(x' - x, y' - y, z' - z) u(x', y', z') dx' dy' dz'.$$

Note that $C = -(\mathcal{L} - c)$. Hence we concentrate on finding suitable operators that agree with C in the bulk. Assuming a separable restricted kernel function similar to (11),

$$C(x, y, z) = X(x)Y(y)Z(z),$$

the operators \mathcal{M}_N and \mathcal{M}_D in 3D defined below enforce pure Neumann and Dirichlet BC and simultaneously agree with the operator \mathcal{L} in the bulk. The construction process is an extension of the 2D case:

$$\begin{aligned}
\mathcal{C} &= \mathcal{C}I_{x',y',z'} = \mathcal{X}I_{x'} \mathcal{Y}I_{y'} \mathcal{Z}I_{z'} \\
&= \mathcal{X}(P_{e,x'} + P_{o,x'}) \mathcal{Y}(P_{e,y'} + P_{o,y'}) \mathcal{Z}(P_{e,z'} + P_{o,z'}) \\
&= (\mathcal{X}P_{e,x'} + \mathcal{X}P_{o,x'}) (\mathcal{Y}P_{e,y'} + \mathcal{Y}P_{o,y'}) (\mathcal{Z}P_{e,z'} + \mathcal{Z}P_{o,z'}) \\
&= (\mathcal{X}_p P_{e,x'} + \mathcal{X}_a P_{o,x'}) (\mathcal{Y}_p P_{e,y'} + \mathcal{Y}_a P_{o,y'}) (\mathcal{Z}_p P_{e,z'} + \mathcal{Z}_a P_{o,z'}) =: -(\mathcal{M}_N - c) \\
&= (\mathcal{X}_a P_{e,x'} + \mathcal{X}_p P_{o,x'}) (\mathcal{Y}_a P_{e,y'} + \mathcal{Y}_p P_{o,y'}) (\mathcal{Z}_a P_{e,y'} + \mathcal{Z}_p P_{o,z'}) =: -(\mathcal{M}_D - c).
\end{aligned}$$

The operators that enforce mixed Neumann and Dirichlet BC can be constructed in a similar fashion to the operators given in (55) and (56). The extension to arbitrary dimension can be performed by the same line of argument.

Numerical Experiments

We numerically solve the following nonlocal wave equation:

$$\begin{aligned}
u_{tt}(\mathbf{x}, t) + \mathcal{M}_{BCU}(\mathbf{x}, t) &= b(\mathbf{x}, t), \quad (\mathbf{x}, t) \in \Omega \times [0, T], \\
u(\mathbf{x}, 0) &= u_0(\mathbf{x}), \\
u_t(\mathbf{x}, 0) &= 0
\end{aligned} \tag{57}$$

by employing the governing operators \mathcal{M}_N and \mathcal{M}_D in 1D, i.e., $BC \in \{N, D\}$, and the operators \mathcal{M}_N , \mathcal{M}_D , $\mathcal{M}_{ND,ND}$, and $\mathcal{M}_{N,DN}$ in 2D, i.e., $BC \in \{N, D, (ND, ND), (N, DN)\}$, with discontinuous and continuous initial displacement $u_0(\mathbf{x})$; see the definition of the governing operators in Theorem 1. For the discretization of the 1D problem, we use the Galerkin projection method with piecewise polynomials. For implementation details and theoretical construction, see Aksoylu et al. (2017a). Note that, for all time, BC are satisfied; see Fig. 3.

In 1D, as far as the boundary behavior goes, in nonlocal problems, we observe a similar wave reflection pattern from the boundary as in classical problems. In the classical case, we see that the Neumann and the Dirichlet BC create reflections of same and opposite signs, respectively; for the Neumann BC, see Fig. 4. A parallel behavior is observed for the nonlocal Neumann and Dirichlet cases; see Fig. 3.

For the discretization of the 2D problem, we use the Nyström method with the quadrature chosen as the trapezoidal rule (Fig. 5). For implementation details, see Aksoylu et al. (Submitted). We depict the solutions to the nonlocal wave equation domain with homogeneous pure Neumann, pure Dirichlet, and mixed Neumann-Dirichlet with vanishing initial velocity and discontinuous initial displacement; see Figs. 6, 7, 8, and 9. Also, for continuous initial displacement, see Fig. 10. The initial solutions are depicted in Fig. 5. Notice that, for all time, local BC are clearly satisfied. Furthermore, for pure Neumann problem, we have numerically verified that $\iint_{\Omega} u(\mathbf{x}, t) d\mathbf{x}$ remains constant for all t . This is in agreement with the physical implication that homogeneous Neumann BC model insulated boundaries.

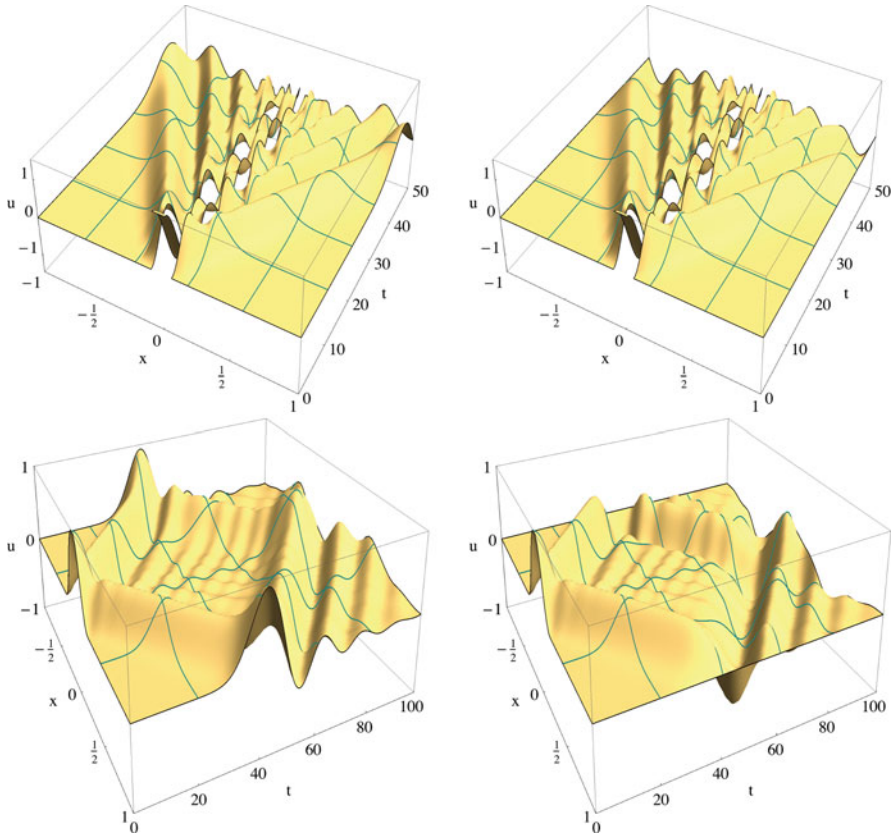


Fig. 3 Solution to the nonlocal wave equation on a 1D domain with discontinuous (top) and continuous (bottom) initial solution and vanishing initial velocity with Neumann (left) boundary condition using the governing operator \mathcal{M}_N and Dirichlet (right) boundary condition using the governing operator \mathcal{M}_D . Note that, for all time, BC are satisfied and discontinuities remain stationary

The Treatment of General Nonlocal Problems Using Functional Calculus

Our main tool that allows us to incorporate local BC into nonlocal operators is *functional calculus*. More precisely, the novel governing operators are obtained by employing the functional calculus of self-adjoint operators, i.e., by replacing the classical governing operator A by a suitable function of A , $f(A)$. We call f the *regulating function*. Since classical BC is an integral part of the classical operator, these BC are automatically inherited by $f(A)$. One advantage of our approach is that every symmetry that commutes with A also commutes with $f(A)$. As a result,

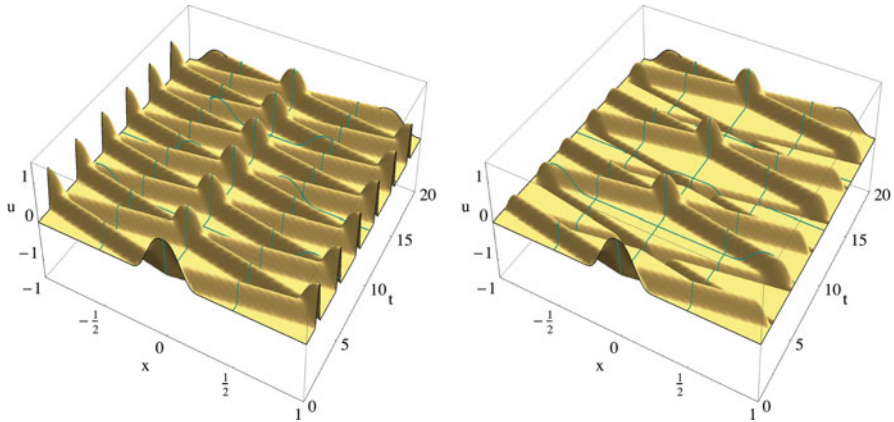


Fig. 4 Solution to the classical wave equation with Neumann (left) and Dirichlet (right) boundary conditions with the same continuous initial displacement as in Fig. 3 and vanishing initial velocity

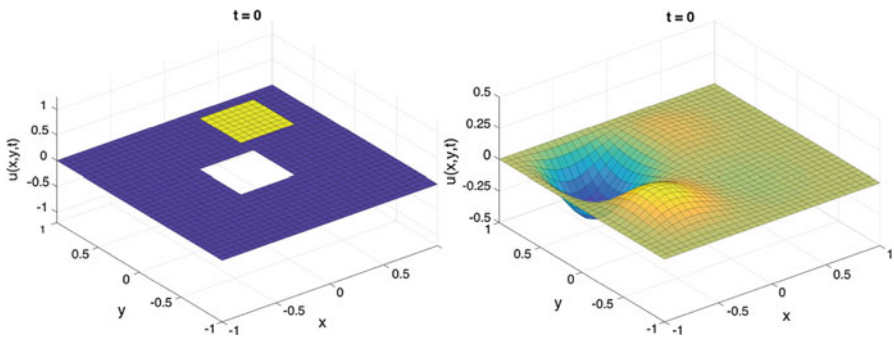


Fig. 5 Initial solutions to the nonlocal wave equation on a 2D domain with discontinuous (left) and continuous (right) initial solutions

required invariance with respect to classical symmetries such as translation, rotation, and so forth is preserved.

We illustrate the benefit of functional calculus, for instance, by comparing the Laplace operator Δ to the biharmonic operator Δ^2 . By simply inspecting how the biharmonic operator is connected to the Laplace operator, one can guess that the regulating function would be $f(\lambda) = \lambda^2$. Before making a rigorous connection, one has to prescribe the BC for each operator. We choose the Laplace operator with homogeneous Dirichlet and Neumann BC and compare it to the biharmonic operator with simply supported (SS) and Cahn-Hilliard (CH) type BC for plate vibration utilizing the weak formulation of the following eigenvalue problems where the BC used are precisely the following:

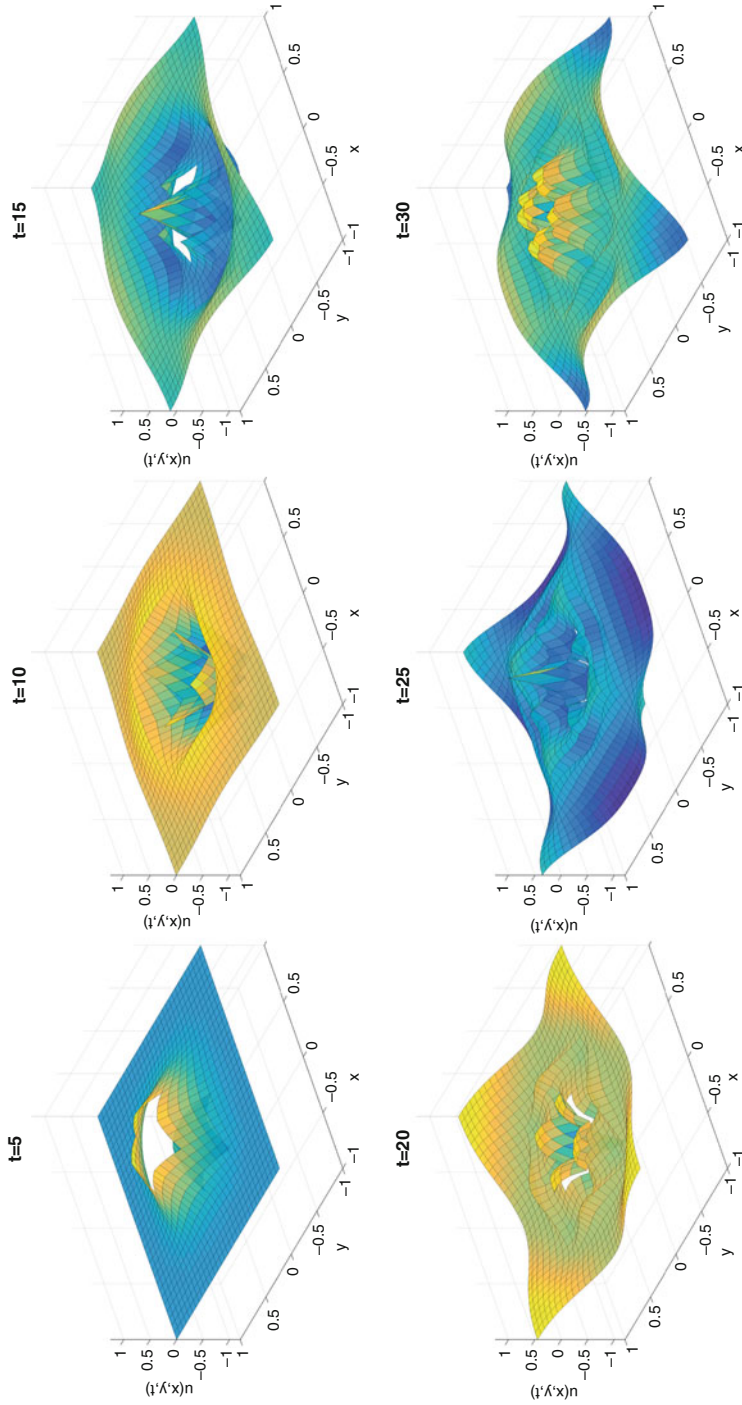


Fig. 6 Solution to the nonlocal wave equation on a 2D domain with discontinuous initial solution using the governing operator \mathcal{M}_N

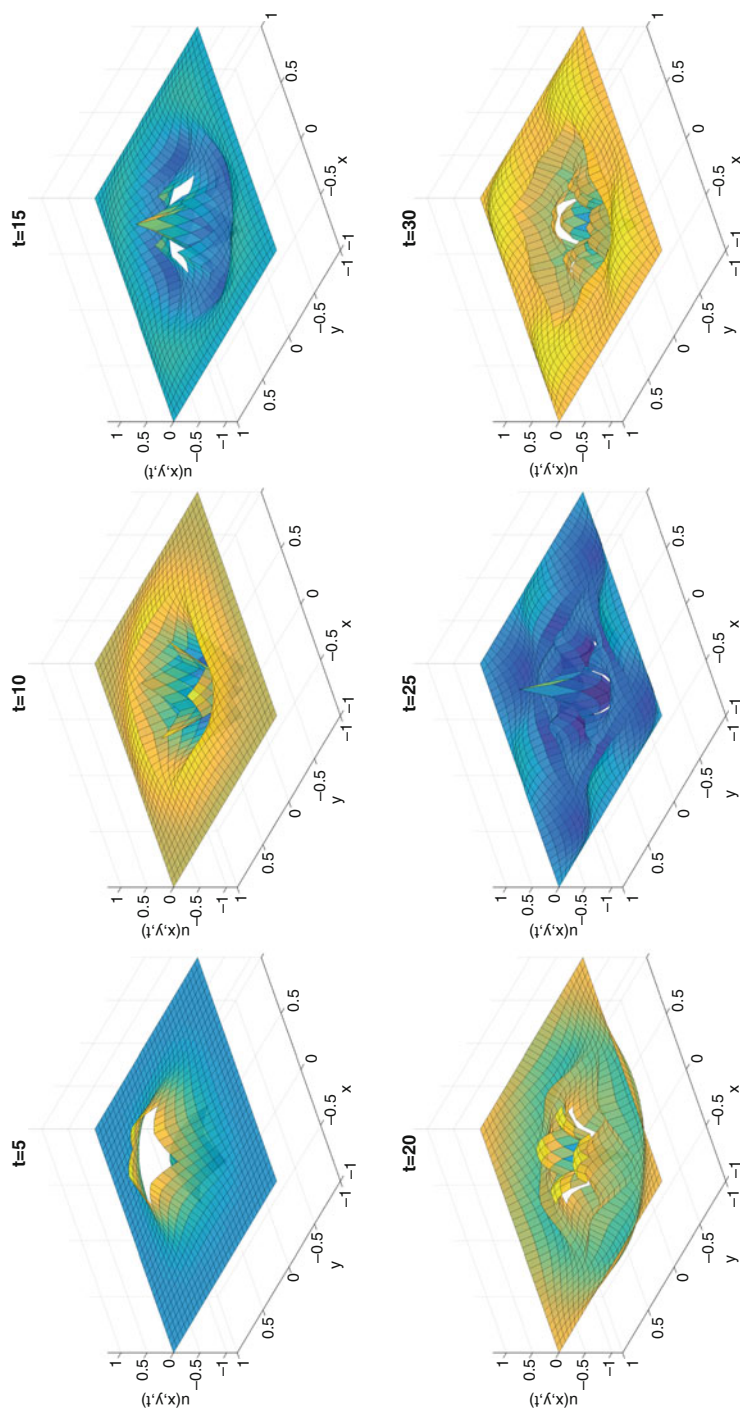


Fig. 7 Solution to the nonlocal wave equation on a 2D domain with discontinuous initial solution using the governing operator \mathcal{M}_D

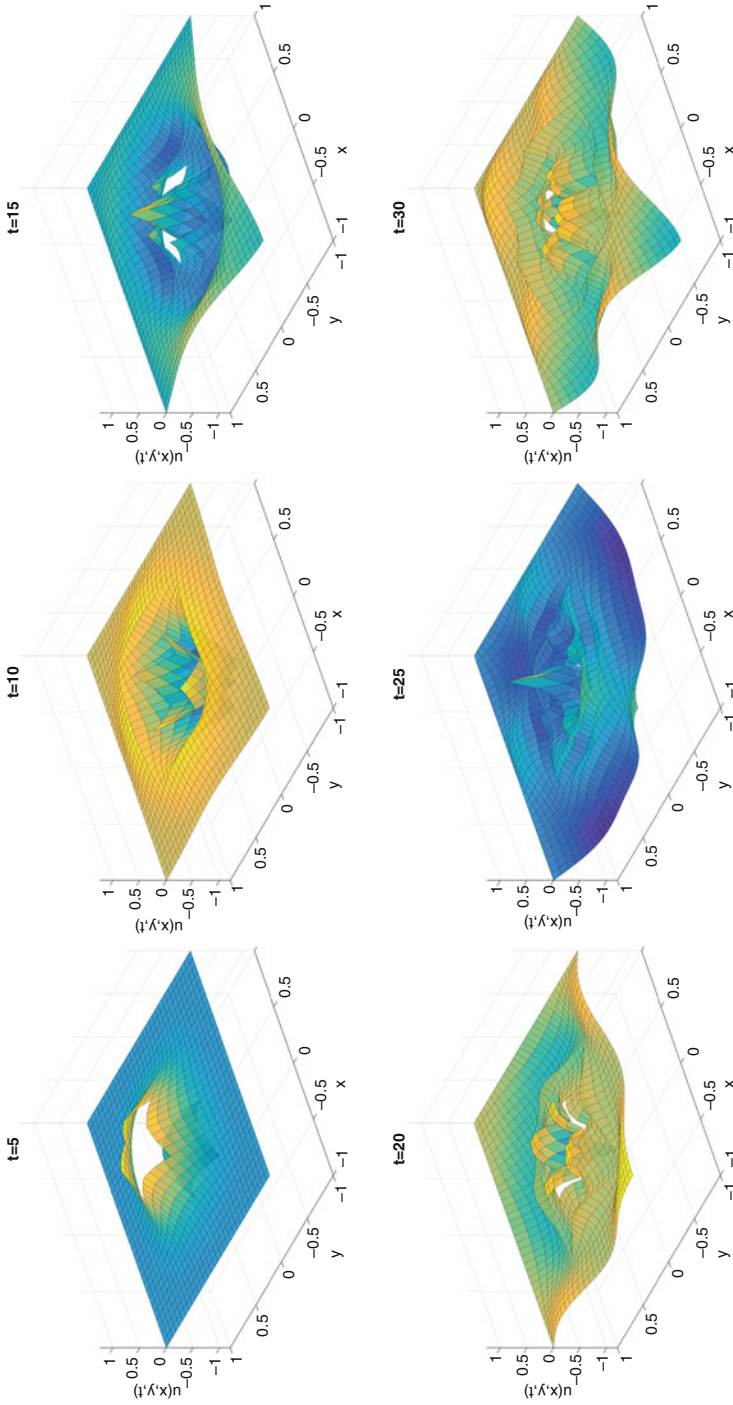


Fig. 8 Solution to the nonlocal wave equation on a 2D domain with discontinuous initial solution using the governing operator $\mathcal{M}_{ND,ND}$

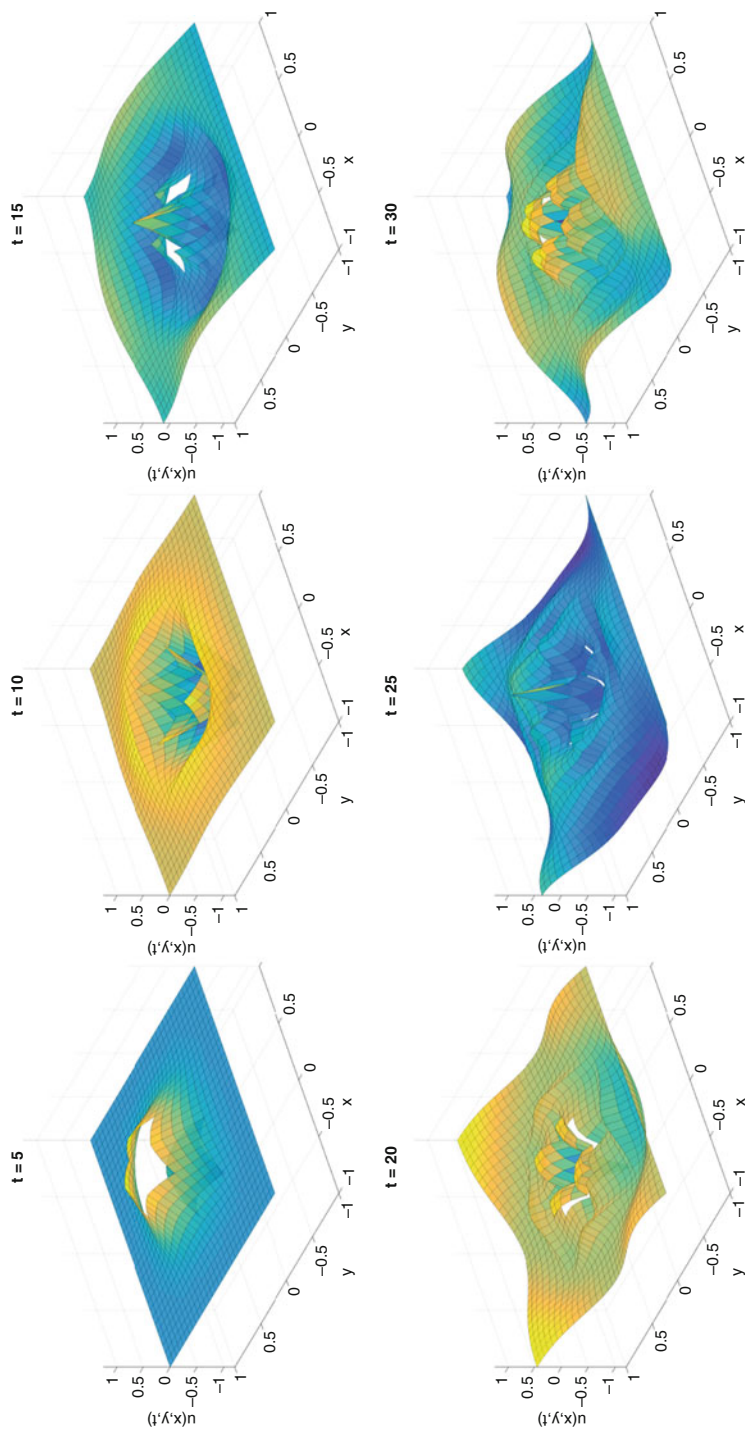


Fig. 9 Solution to the nonlocal wave equation on a 2D domain with discontinuous initial solution using the governing operator $\mathcal{M}_{N, DN}$

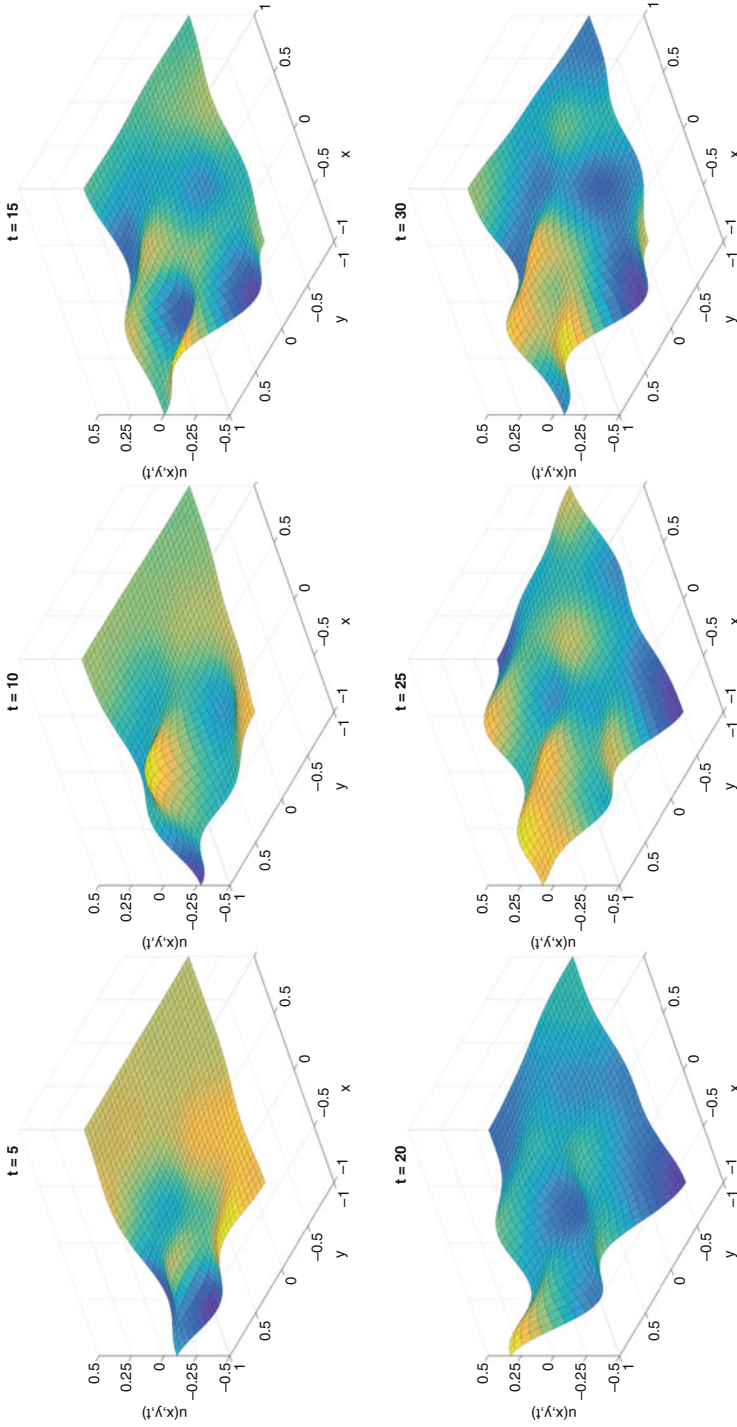


Fig. 10 Solution to the nonlocal wave equation on a 2D domain with continuous initial solution using the governing operator \mathcal{M}_N

$$\text{SS-BC: } u = \Delta u = 0 \text{ on } \partial\Omega, \quad \text{CH-BC: } \frac{\partial u}{\partial n} = \frac{\partial \Delta u}{\partial n} = 0 \text{ on } \partial\Omega.$$

$$L(u, v) := \int_{\Omega} \nabla u \cdot \nabla v \, dx = \lambda R(u, v) := \lambda \int_{\Omega} uv \, dx, \quad v \in V_{L,D} \text{ or } V_{L,N}$$

$$B(u, v) := \int_{\Omega} \nabla^2 u : \nabla^2 v \, dx = \lambda R(u, v), \quad v \in V_{B,SS} \text{ or } V_{B,CH},$$

$$V_{L,D} := H_0^1(\Omega), \quad V_{L,N} := \left\{ v \in H^1(\Omega) : \frac{\partial v}{\partial n} = 0 \right\},$$

$$V_{B,SS} := H_0^1(\Omega) \cap H^2(\Omega), \quad V_{B,CH} := \left\{ v \in H^2(\Omega) : \frac{\partial v}{\partial n} = 0 \right\}.$$

Indeed, the eigenvalues of the biharmonic operator with SS-BC and CH-BC are the squares of those of the Laplace operator with Dirichlet and Neumann BC, respectively. Furthermore, the eigenfunctions are identical for Dirichlet and Neumann BC with SS and CH, respectively. We have provided this example as a proof of concept and, hence, chosen the BC carefully to establish the connection. One may not obtain such connection with arbitrary BC.

The convolution operators in (12) in the form of integrals are derived from their (original) series representation. We defined generalized convolution operators in Aksoylu et al. (2017a,b) in the following series form:

$$C_{BC}u(x) := \sum_k \langle e_k^{BC} | C \rangle \langle e_k^{BC} | u \rangle e_k^{BC}, \tag{58}$$

where $BC = p, a$ and $\langle \cdot | \cdot \rangle$ denotes the inner product in $L_C^2(\Omega)$ and is defined by

$$\langle e_k^{BC} | u \rangle := \int_{\Omega} (e_k^{BC})^*(x') u(x') \, dx'.$$

In addition, $(e_k^{BC})_k$ is chosen to be a basis associated with a multiple of the Laplace operator with appropriate BC, which we call as the classical operator and denote by Δ_{BC} . The spectrum of Δ_{BC} with classical BC such as periodic, antiperiodic, Neumann, and Dirichlet is purely discrete. Furthermore, we can explicitly calculate the eigenfunctions e_k^{BC} corresponding to each BC. These eigenfunctions form a Hilbert (complete and orthonormal) basis for $L_C^2(\Omega)$ through which the generalized convolution operator is defined. The main reason why we discuss Δ_{BC} is the fact that the governing operator (1) turns out to be a function of Δ_{BC} (Aksoylu et al. 2017a,b; Beyer et al. 2016). Since the classical operator A_{BC} is defined through local BC, the eigenfunctions inherit this information. This observation opened a gateway to incorporate local BC to nonlocal theories on bounded domains (Aksoylu et al. 2017b).

The normalized eigenfunctions of the classical operators are as follows:

$$e_k^p(x) := \frac{1}{\sqrt{2}} e^{i\pi kx}, \quad k \in \mathbb{N}, \quad e_k^a(x) := \frac{1}{\sqrt{2}} e^{i\pi(k+\frac{1}{2})x}, \quad k \in \mathbb{N}, \tag{59}$$

$$e_k^N(x) := \begin{cases} \frac{1}{\sqrt{2}}, & k = 0, \\ \cos\left(\frac{k\pi}{2}(x+1)\right), & k \in \mathbb{N}^*, \end{cases} \quad e_k^D(x) := \sin\left(\frac{k\pi}{2}(x+1)\right), \quad k \in \mathbb{N}^*.$$

Plugging the eigenfunctions in (59) into (58) and after hefty calculation, we proved that the operators C_p and C_a have integral representations given (12). For more details, see Aksoylu et al. (2017a).

Next, we present the steps how to apply functional calculus (FC). We denote a nonlocal operator by NL and its local counterpart by A . Note that both nonlocal diffusion and PD operators are defined initially on \mathbb{R}^d and contain convolution. The size of nonlocality is determined by the parameter δ which is encoded in the kernel function.

FC-1. Apply limit to the horizon parameter, i.e., $\delta \rightarrow 0$, to identify a local counterpart A of NL .

FC-2. Apply the Fourier transform to “diagonalize” NL and A to obtain the corresponding spectra.

FC-3. Read off the regulating function f by comparing the spectra of NL and A . Spectra on \mathbb{R}^d are continuous.

We apply the above steps to the concrete example of nonlocal diffusion on \mathbb{R}^d where the classical operator A is the Laplace operator $-\Delta : W^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$.

$$NLu(\mathbf{x}, t) = f(A)u(\mathbf{x}, t) := \left(\int_{\mathbb{R}^d} C(\mathbf{x}') d\mathbf{x}' \right) u(\mathbf{x}, t) - \int_{\mathbb{R}^d} C(\mathbf{x}' - \mathbf{x}) u(\mathbf{x}', t) d\mathbf{x}'. \tag{60}$$

We connect the nonlocal operator to A through Fourier transforms. Let $F_1 : L^1(\mathbb{R}^d) \rightarrow C_\infty(\mathbb{R}^d)$, $F_2 : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$ be the Fourier transforms and the kernel function $C \in L^1(\mathbb{R}^d)$ be even:

$$A = F_2^{-1} \circ T_\lambda \circ F_2, \tag{61}$$

$$f(A) = F_2^{-1} \circ T_{f(\lambda)} \circ F_2,$$

where $T_{h(\lambda)}$ denotes the maximal multiplication operator by $h(\lambda)$. Then, we directly diagonalize $f(A)$ by using the expression given in (60):

$$f(A) = F_2^{-1} \circ T_{F_1 C(0) - F_1 C(\lambda)} \circ F_2. \tag{62}$$

Therefore, a comparison of (61) and (62) yields

$$f(\lambda) = F_1 C(0) - F_1 C(\lambda).$$

We explicitly identify to which function of the classical operator the nonlocal operator corresponds:

$$f(\lambda) = F_1 C(0) - F_1 C(\lambda) = \langle 1|C \rangle - \langle e_\lambda|C \rangle, \quad \langle e_\lambda|C \rangle := \int_{\mathbb{R}^d} e_\lambda(\mathbf{x}')^* C(\mathbf{x}') d\mathbf{x}',$$

where the spectral value of the classical operator $\lambda \in [0, \infty)$. Now, we extend the construction on \mathbb{R}^d to a bounded domain Ω .

FC-4. Restrict A to Ω with a prescribed BC. Denote the new operator by A_{BC} .

Spectrum of A_{BC} , $\sigma(A_{\text{BC}})$ is now discrete. Find the eigenfunctions of A_{BC} .

FC-5. Define a generalized convolution as in (58) by using eigenfunctions of A_{BC} .

FC-6. Rewrite (recycle) the regulating function with discrete spectrum.

$$f_{\text{BC}} : \sigma(A_{\text{BC}}) \rightarrow \mathbb{R} \quad f_{\text{BC}}(\lambda_k^{\text{BC}}) = \langle 1|C \rangle - \langle e_k^{\text{BC}}|C \rangle, \quad \text{BC} \in \{\text{p}, \text{a}\}. \quad (63)$$

FC-7. Construct $f_{\text{BC}}(A_{\text{BC}})$ using the spectral theorem. Namely, for $u = \sum_k \langle e_k^{\text{BC}}|u \rangle e_k^{\text{BC}}$, we have

$$f_{\text{BC}}(A_{\text{BC}})u = \sum_k f_{\text{BC}}(\lambda_k^{\text{BC}}) \langle e_k^{\text{BC}}|u \rangle e_k^{\text{BC}}. \quad (64)$$

FC-8. Find a computationally feasible expression of $f_{\text{BC}}(A_{\text{BC}})$ such as an integral representation.

Now, we show how we use the **FC** steps to construct the governing operators \mathcal{M}_{BC} , $\text{BC} \in \{\text{p}, \text{a}\}$ in 1D. Namely, we want to verify $f_{\text{BC}}(A_{\text{BC}})u = \mathcal{M}_{\text{BC}}u$.

Using (63) and (64), we have the following:

$$\begin{aligned} f_{\text{BC}}(A_{\text{BC}})u &= \sum_k [\langle 1|C \rangle - \langle e_k^{\text{BC}}|C \rangle] \langle e_k^{\text{BC}}|u \rangle e_k^{\text{BC}} \\ &= \langle 1|C \rangle \sum_k \langle e_k^{\text{BC}}|u \rangle e_k^{\text{BC}} - \sum_k \langle e_k^{\text{BC}}|C \rangle \langle e_k^{\text{BC}}|u \rangle e_k^{\text{BC}} \\ &= cu - \mathcal{C}_{\text{BC}}u \\ &= \mathcal{M}_{\text{BC}}, \quad \text{BC} \in \{\text{p}, \text{a}\}. \end{aligned}$$

Expressing the regulating function for the case of $\text{BC} \in \{\text{N}, \text{D}\}$ requires nontrivial manipulation of series and is more involved than the case of $\text{BC} \in \{\text{N}, \text{D}\}$. We simply report them here:

$$f_N : \sigma(A_N) \rightarrow \mathbb{R}, \quad f_N(\lambda_k^N) = \langle 1|C \rangle - \begin{cases} \langle e_{k/2}^D|C \rangle & \text{if } k \in \mathbb{N} \text{ is even,} \\ \langle e_{(k-1)/2}^A|C \rangle & \text{if } k \in \mathbb{N} \text{ is odd.} \end{cases}$$

$$f_D : \sigma(A_D) \rightarrow \mathbb{R}, \quad f_D(\lambda_k^D) = \langle 1|C \rangle - \begin{cases} \langle e_{k/2}^D|C \rangle & \text{if } k \in \mathbb{N}^* \text{ is even,} \\ \langle e_{(k-1)/2}^A|C \rangle & \text{if } k \in \mathbb{N} \text{ is odd.} \end{cases}$$

In Aksoylu et al. (2017a), we showed that

$$f_N(A_N)u = (c - C_p P_e - C_a P_o)u = \mathcal{M}_N u$$

$$f_D(A_D)u = (c - C_a P_e - C_p P_o)u = \mathcal{M}_D u.$$

The operators \mathcal{M}_N and \mathcal{M}_D were used as governing operator in (57) to perform the numerical experiments in 1D.

Remark 7. Fractional diffusion and fractional PDEs also fall into the class of nonlocal problems; see some of the recent developments (Andreu-Vaillo et al. 2010; Caffarelli et al. 2007; Di Nezza et al. 2012; Nochetto et al. 2015). There is a fundamental difference between these operators and ours: our governing operators are bounded. Note that the regulating function in (63) is bounded and that is why the application of the spectral theorem in (64) is valid. Since our ultimate goal is to capture discontinuities or cracks, we are mainly interested in bounded governing operators. Fractional operators become unbounded for such discontinuities, and hence, we exclude them from our discussion.

Conclusion

We presented novel governing operators in arbitrary dimension for nonlocal diffusion. The operators agree with the original PD operator in the bulk of the domain and simultaneously enforce local BC. We presented methodically how to verify the BC by using a change in the order of integration. We presented different types of BC in 2D which include pure and mixed combinations of Neumann and Dirichlet BC. We presented numerical experiments for the nonlocal wave equation. We verified that the novel operators enforce local BC for all time. We also observed that the property we proved for 1D, namely, discontinuities remain stationary, also holds for 2D.

Our ongoing work aims to extend the novel operators to vector-valued problems which will allow the extension of PD to applications that require local BC. Furthermore, we hope that our novel approach potentially will avoid altogether the surface effects seen in PD.

Acknowledgements Burak Aksoylu was supported in part by the European Commission Marie Curie Career Integration 293978 grant, and Scientific and Technological Research Council of Turkey (TÜBİTAK) MFAG 115F473 grant.

References

- B. Aksoylu, H.R. Beyer, F. Celiker, Application and implementation of incorporating local boundary conditions into nonlocal problems. *Numer. Funct. Anal. Optim.* **38**(9), 1077–1114 (2017a). <https://doi.org/10.1080/01630563.2017.1320674>
- B. Aksoylu, H.R. Beyer, F. Celiker, Theoretical foundations of incorporating local boundary conditions into nonlocal problems. *Rep. Math. Phys.* **40**(1), 39–71 (2017b). [https://doi.org/10.1016/S0034-4877\(17\)30061-7](https://doi.org/10.1016/S0034-4877(17)30061-7)
- B. Aksoylu, F. Celiker, Comparison of nonlocal operators utilizing perturbation analysis, in *Numerical Mathematics and Advanced Applications ENUMATH 2015*, vol. 112, ed. by B.K. et al. Lecture Notes in Computational Science and Engineering (Springer, 2016), pp. 589–606. https://doi.org/10.1007/978-3-319-39929-4_57
- B. Aksoylu, F. Celiker, Nonlocal problems with local Dirichlet and Neumann boundary conditions. *J. Mech. Mater. Struct.* **12**(4), 425–437 (2017). <https://doi.org/10.2140/jomms.2017.12.425>
- B. Aksoylu, F. Celiker, O. Kilicer, Nonlocal problems with local boundary conditions in higher dimensions (Submitted)
- B. Aksoylu, A. Kaya, Conditioning and error analysis of nonlocal problems with local boundary conditions. *J. Comput. Appl. Math.* **335**, 1–19 (2018). <https://doi.org/10.1016/j.cam.2017.11.023>
- B. Aksoylu, T. Mengesha, Results on nonlocal boundary value problems. *Numer. Funct. Anal. Optim.* **31**(12), 1301–1317 (2010)
- B. Aksoylu, M.L. Parks, Variational theory and domain decomposition for nonlocal problems. *Appl. Math. Comput.* **217**, 6498–6515 (2011). <https://doi.org/10.1016/j.amc.2011.01.027>
- B. Aksoylu, Z. Unlu, Conditioning analysis of nonlocal integral operators in fractional Sobolev spaces. *SIAM J. Numer. Anal.* **52**(2), 653–677 (2014)
- F. Andreu-Vaillio, J.M. Mazon, J.D. Rossi, J. Toledo-Melero, *Nonlocal Diffusion Problems*. Mathematical Surveys and Monographs, vol. 165. (American Mathematical Society and Real Socied Matematica Espanola, 2010)
- H.R. Beyer, B. Aksoylu, F. Celiker, On a class of nonlocal wave equations from applications. *J. Math. Phys.* **57**(6), 062902 (2016). <https://doi.org/10.1063/1.4953252>. Eid:062902
- L. Caffarelli, L. Silvestre, An extension problem related to the fractional Laplacian. *Commun. Part. Diff. Eqs.* **32**, 1245–1260 (2007)
- E. Di Nezza, G. Palatucci, E. Valdinoci, Hitchhiker’s guide to fractional Sobolev spaces. *Bull. Sci. Math.* **136**(5), 521–573 (2012)
- Q. Du, M. Gunzburger, R.B. Lehoucq, K. Zhou, Analysis and approximation of nonlocal diffusion problems with volume constraints. *SIAM Rev.* **54**, 667–696 (2012)
- Q. Du, R. Lipton, Peridynamics, fracture, and nonlocal continuum models. *SIAM News* **47**(3) (2014)
- E. Emmrich, O. Weckner, The peridynamic equation and its spatial discretization. *Math. Model. Anal.* **12**(1), 17–27 (2007)
- G. Gilboa, S. Osher, Nonlocal operators with applications to image processing. *Multiscale Model. Simul.* **7**(3), 1005–1028 (2008)
- B. Kilic, Peridynamic theory for progressive failure prediction in homogeneous and heterogeneous materials. Ph.D. thesis, Department of Aerospace and Mechanical Engineering, University of Arizona, Tucson (2008)
- E. Madenci, E. Oterkus, *Peridynamic Theory and Its Applications* (Springer, New York/Heidelberg/Dordrecht/London, 2014). <https://doi.org/10.1007/978-1-4614-8465-3>

- T. Mengesha, Q. Du, Analysis of a scalar peridynamic model for sign changing kernel. *Disc. Cont. Dyn. Sys. B* **18**, 1415–1437 (2013)
- J.A. Mitchell, S.A. Silling, D.J. Littlewood, A position-aware linear solid constitutive model for peridynamics. *J. Mech. Mater. Struct.* **10**(5), 539–557 (2015)
- R.H. Nochetto, E. Otárola, A.J. Salgado, A PDE approach to fractional diffusion in general domains: a priori error analysis. *Found. Comput. Math.* **15**, 733–791 (2015)
- P. Seleson, M. Gunzburger, M.L. Parks, Interface problems in nonlocal diffusion and sharp transitions between local and nonlocal domains. *Comput. Methods Appl. Mech. Eng.* **266**, 185–204 (2013)
- P. Seleson, M.L. Parks, On the role of the influence function in the peridynamic theory. *Internat. J. Multiscale Comput. Eng.* **9**(6), 689–706 (2011)
- S. Silling, Reformulation of elasticity theory for discontinuities and long-range forces. *J. Mech. Phys. Solids* **48**, 175–209 (2000)
- S. Silling, R.B. Lehoucq, Peridynamic theory of solid mechanics. *Adv. Appl. Mech.* **44**, 73–168 (2010)
- S.A. Silling, M. Zimmermann, R. Abeyaratne, Deformation of a peridynamic bar. *J. Elast.* **73**, 173–190 (2003)
- X. Tian, Q. Du, Analysis and comparison of different approximations to nonlocal diffusion and linear peridynamic equations. *SIAM J. Numer. Anal.* **51**(6), 3458–3482 (2013)
- O. Weckner, R. Abeyaratne, The effect of long-range forces on the dynamics of a bar. *J. Mech. Phys. Solids* **53**, 705–728 (2005)



Peridynamics and Nonlocal Diffusion Models: Fast Numerical Methods

40

Hong Wang

Contents

Introduction	1332
Nonlocal Diffusion and Peridynamic Models	1334
A Nonlocal Diffusion Model	1334
A Bond-Based Linear Peridynamic Model	1335
Numerical Issues and Computational Complexity	1336
Fast Numerical Methods for a One-Dimensional Peridynamic Model	1337
A Fast Galerkin Finite Element Method on a Uniform Partition	1338
A Fast Collocation Method on a Graded Mesh	1340
A Fast Collocation Method for the Nonlocal Diffusion Model (1)	1342
A Fast Collocation Method for the Peridynamic Model (3)	1346
Concluding Remarks	1350
References	1351

Abstract

We outline the recent developments of fast numerical methods for linear nonlocal diffusion and peridynamic models in one and two space dimensions. We show how the analysis was carried out to take full advantage of the structure of the stiffness matrices of the numerical methods in its storage, evaluation, and assembly and in the efficient solution of the corresponding numerical schemes. This significantly reduces the computational complexity and storage of the numerical methods over conventional ones, without using any lossy compression. For instance, we would use the same numerical quadratures for conventional methods to evaluate the singular integrals in the stiffness matrices, except that we only need to evaluate $O(N)$ of them instead of $O(N^2)$ of them. Numerical results are presented to show the utility of these fast methods.

H. Wang (✉)

Department of Mathematics, University of South Carolina, Columbia, SC, USA

e-mail: hwang@math.sc.edu

Keywords

Peridynamics · Nonlocal diffusion model · Fast numerical method · Toplitz matrix

Introduction

In the last couple of decades, many nonlocal models have been developed, which are emerging as powerful tools for modeling challenging phenomena in many disciplines and applications, including problems involving anomalous transport and long-range time memory or spatial interactions. For instance, in the scenario of diffusion processes, Fick first set up the diffusion equation at a macroscopic scale in a deterministic manner during his study on how nutrients travel through membranes in living organisms. But Fick's approach was phenomenological. It was Einstein who derived the diffusion equation from first principle as part of his work on Brownian motion. Alternatively, Pearson modeled diffusion process via a stochastic formulation at microscopic scale in terms of a random walk, leading to the stochastic differential equation driven by a Brownian motion (Meerschaert and Sikorskii 2011; Metzler and Klafter 2000; Øksendal 2010; Podlubny 1999).

The common assumptions between Einstein's explanation of diffusion and Pearson's random walk are (i) the existence of a mean free path and (ii) the existence of a mean waiting time. Under these assumptions, the central limit theorem concludes that the underlying stochastic process that describes the particle jumps is a Brownian motion in the Pearson's approach. Equivalently, the corresponding probability density function describing the stochastic process satisfies the Fickian diffusion equation as the Fokker-Planck equation of the Brownian motion in the Einstein's approach (Meerschaert and Sikorskii 2011). Note that the assumptions (i) and (ii), which lead to a stochastic process of Brownian motion, virtually hold for transport processes in homogeneous media. In highly heterogeneous media with faults and fractures, the underlying particle movements in the transport processes may experience long jumps and so will have large deviation from Brownian motion but converge to a Lévy process. Consequently, the corresponding probability density function satisfies a fractional partial differential equation (PDE) as its Fokker-Planck equation (Meerschaert and Sikorskii 2011). This justifies why a fractional PDE or a nonlocal diffusion model better describes anomalous diffusive processes in heterogeneous media than traditional integer-order PDEs (del-Castillo-Negrete et al. 2004; D'Elia et al. 2014; Du et al. 2012; Meerschaert and Sikorskii 2011; Metzler and Klafter 2000; Podlubny 1999).

Similar phenomena occur in the context of continuum solid mechanics. In this case, the classical theory assumes that all internal forces act locally and yields mathematical models that are expressed in terms of PDEs. These models have difficulties to describe problems with evolving discontinuities, due to its differentiability assumption on displacement fields. The peridynamic theory (Silling 2000; Silling et al. 2007) was proposed as a reformation of the classical theory of continuum solid mechanics and yields nonlocal mathematical formulations that are

based on long-range interactions. Constitutive models in the peridynamic theory depend on finite deformation vectors, instead of deformation gradients in classical constitutive models. Consequently, peridynamic models are particularly suitable for the representation of discontinuities in displacement fields and the description of cracks and their evolution in materials. To date, peridynamic models have been successfully applied in many important applications, including failure and damage in composite laminates (Oterkus et al. 2012), crack propagation and branching (Ha and Bobaru 2011), crack nucleation (Silling et al. 2010; Parks et al. 2008; Seleson et al. 2009), phase transformations in solids (Dayal and Bhattacharya 2006), impact damage (Seleson and Parks 2011; Silling and Askari 2005), polycrystal fracture (Ghajari et al. 2014), crystal plasticity (Sun and Sundararaghavan 2014), damage in concrete (Gerstle et al. 2007), and geomaterial fragmentation (Lai et al. 2015).

Because these nonlocal models provide greatly improved modeling capability than traditional integer-order PDE models do, different numerical methods have been developed for these models with different computational expense, memory requirements, and implementational effort as well as accuracy, convergence, and stability (Emmrich and Weckner 2007). For example, collocation methods and mesh-free methods apply directly to the strong form of the nonlocal models and are relatively simple to implement (Seleson et al. 2016; Seleson and Littlewood 2016; Silling and Askari 2005). On the other hand, finite element discretizations of the nonlocal models are based on Galerkin weak forms and enjoy high convergence rates (Chen and Gunzburger 2011; Du et al. 2013; Oterkus et al. 2012; Tian and Du 2013; Zhou and Du 2010).

However, nonlocal models involve singular integral operators and present numerical difficulties that were not encountered in classical integer-order PDE models: (i) the numerical methods for integer-order PDE models generate sparse stiffness matrices which have a linear storage requirement of $O(N)$ where N is the number of spatial unknowns. The matrix-vector multiplication has computational work of $O(N)$. Hence, the convergence of any Krylov subspace iterative solver depends only on the number of iterations. In contrast, the numerical methods for nonlocal models yield dense stiffness matrices in which their bandwidths increase to infinity as the mesh size decreases to zero. Consequently, the numerical methods for nonlocal models have $O(N^2)$ memory requirement. Direct solvers have been widely used in the numerical simulation of nonlocal models, which have an $O(N^3)$ computational complexity in order to obtain the numerical solutions. On the other hand, the matrix-vector multiplication by the stiffness matrices has $O(N^2)$ computational complexity. (ii) Due to the impact of the singular kernels in the integral operators in nonlocal models, the evaluation and assembly of the stiffness matrices require the evaluation of $O(N^2)$ entries, which can be very expensive. Numerical experiments show that the evaluation and assembly of the coefficient matrix often constitute a very large portion of simulation times! (iii) While they have high-order convergence rates, Galerkin finite element methods double the number of spatial dimensions that need to be discretized, leading to a significantly increased computational work and memory requirement as compared to collocation methods. In summary, the significantly increased computational work and memory requirement of the

nonlocal models over those for the classical integer-order PDE models severely limit the applications of these models, especially for problems in multiple space dimensions.

In this chapter we go over the recent developments of fast and accurate numerical methods for nonlocal models. We do not intend to develop a more accurate numerical discretization or numerical quadrature for these models. Rather, we explore the structure of the stiffness matrices of existing numerical methods in order to significantly reduce the computational work to evaluate and assemble the stiffness matrices and to compute the numerical solutions, as well as the memory requirement to store the stiffness matrices. The idea applies to different numerical methods and numerical quadratures in the literature. In this chapter we take some representative numerical methods as examples to illustrate the developments.

The rest of this chapter is organized as follows. In section “[Nonlocal Diffusion and Peridynamic Models](#)” we present a nonlocal diffusion model and a bond-based peridynamic model and discuss related numerical issues. In section “[Fast Numerical Methods for a One-Dimensional Peridynamic Model](#)” we go over the development of fast Galerkin finite element method and a fast collocation method for a one-dimensional peridynamic model. In section “[A Fast Collocation Method for the Nonlocal Diffusion Model \(1\)](#)” we outline the development of a fast collocation method for a nonlocal diffusion model in two space dimensions. In section “[A Fast Collocation Method for the Peridynamic Model \(3\)](#)” we go over the development of a fast collocation method for a bond-based peridynamic model in two space dimensions. In section “[Concluding Remarks](#)” we draw concluding remarks and discuss the future directions.

Nonlocal Diffusion and Peridynamic Models

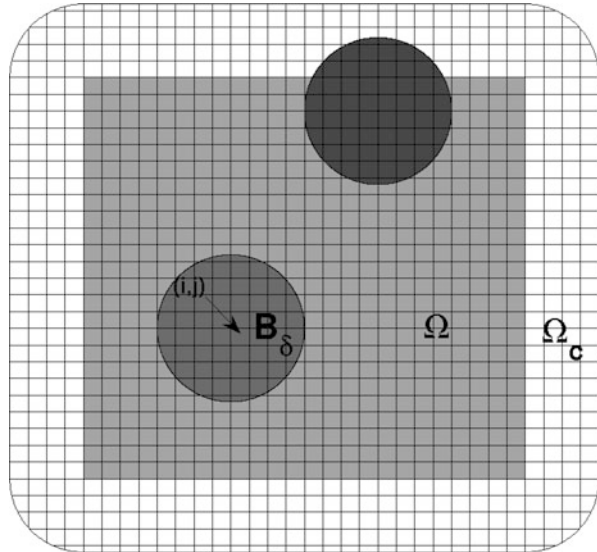
In this section we use the presentative nonlocal diffusion model and peridynamic model in two space dimensions as an example to illustrate the corresponding numerical issues.

A Nonlocal Diffusion Model

A linear nonlocal diffusion model can be expressed as the following integral-differential equation (D’Elia et al. 2014; Du et al. 2012):

$$\begin{aligned}
 u_t(x, y, t) + \int_{B_\delta(x, y)} K(x - x', y - y') (u(x, y, t) - u(x', y', t)) dx' dy' \\
 &= f(x, y, t), \quad (x, y) \in \Omega, \quad t \in (0, T], \\
 u(x, y, 0) &= u_o(x, y), \quad (x, y) \in \Omega, \\
 u(x, y, t) &= g(x, y, t), \quad (x, y) \in \Omega_c, \quad t \in [0, T].
 \end{aligned} \tag{1}$$

Fig. 1 Illustration of a spatial partition on $\Omega \cup \Omega_c$ with a circular horizon



Here Ω is the physical domain in the plane, Ω_c represents a boundary zone of width δ surrounding Ω with $\delta > 0$ being the horizon parameter of the material, u represents the concentration of the solute, u_0 is the prescribed initial condition, and f is the source and sink term, respectively; $K(x, y)$ is the kernel function of the integral operator which is often of the form

$$K(x, y) := \frac{C}{(x^2 + y^2)^{1+\gamma}} \tag{2}$$

with C being a normalization constant; $B_\delta(x, y)$ represents the material horizon which is often chosen to be a disk centered at (x, y) with radius δ (cf. Fig. 1).

We note that in the nonlocal diffusion model (1), the boundary condition is “volume constrained” that is imposed on a two-dimensional boundary zone Ω_c . This is in contrast to the boundary conditions for classical integer-order PDE models that are imposed on the boundary $\partial\Omega$ of the physical domain Ω . It is clear that the integral operator in (1) can be viewed as a truncated version of the fractional Laplacian operator $(-\Delta)^\gamma$, which is the infinitesimal generator of a Feller semigroup that is generated by a Lévy process (Applebaum 2009; Meerschaert and Sikorskii 2011).

A Bond-Based Linear Peridynamic Model

In a bond-based linear peridynamic model for modeling a proportional, microelastic material in a reference domain in the plane, the equation of motion at any point

(x, y) in the reference configuration Ω at any time $t \in (0, T]$ can be expressed as (Silling 2000)

$$\begin{aligned} \rho \mathbf{u}_{tt}(x, y, t) + \int_{B_\delta(x, y)} \mathbf{K}(x - x', y - y')(\mathbf{u}(x, y, t) - \mathbf{u}(x', y', t)) dx' dy' \\ = \mathbf{f}(x, y, t), \quad (x, y) \in \Omega, \quad t \in (0, T], \\ \mathbf{u}(x, y, 0) = \mathbf{u}_o(x, y), \quad (x, y) \in \Omega, \\ \mathbf{u}(x, y, t) = \mathbf{g}(x, y, t), \quad (x, y) \in \Omega_c, \quad t \in [0, T]. \end{aligned} \tag{3}$$

Here $\mathbf{u}(x, y, t) = [v(x, y, t), w(x, y, t)]^T$ represents the displacement vector field with $v(x, y, t)$ and $w(x, y, t)$ being its components in the x and y directions, respectively; similarly, $\mathbf{f}(x, y, t) = [f^v(x, y, t), f^w(x, y, t)]^T$ represents an external force density field, and $\mathbf{g}(x, y, t) = [g^v(x, y, t), g^w(x, y, t)]^T$ is the prescribed nonlocal boundary data imposed on the domain Ω_c . ρ is the mass density in the reference configuration. The micromodulus function for the material or the kernel of the integral operator \mathbf{K} is of the form

$$\mathbf{K}(x, y) = \frac{c}{(x^2 + y^2)^{3/2}} \begin{bmatrix} x^2 & xy \\ xy & y^2 \end{bmatrix} \tag{4}$$

where c is a constant that depends on the material properties and space dimensions.

Numerical Issues and Computational Complexity

The mathematical analysis for the nonlocal and peridynamic models and their corresponding numerical analysis can be found, e.g., in Chen and Gunzburger (2011), D’Elia et al. (2014), Du et al. (2012, 2013), Tian and Du (2013), Zhang et al. (2016), and Zhou and Du (2010). In particular, the following error estimate which was proved in Zhou and Du (2010) for the linear finite element method for a steady-state nonlocal diffusion or peridynamic model states that for any $0 < \varepsilon \ll 1$, there exists a positive constant $C = C(\varepsilon)$ which is independent of h such that

$$\|u - u^h\|_{L^2} \leq Ch^{2-\varepsilon} \|u\|_{H^2}. \tag{5}$$

The estimate looks similar to that for the finite element method for classical integer-order PDEs in which the stiffness matrix is sparse. In this case, the stiffness matrix can be stored in $O(N)$ memory and each matrix-vector multiplication (so each Krylov subspace iteration) can be carried out in $O(N)$ computations.

However, for the nonlocal diffusion model (1) and peridynamic model (3), the singular integral operator is defined on the material horizon $B_\delta(x, y)$, which yields a dense stiffness matrix that has $O(N)$ nonzero entries at each row, as the mesh size

h tends to zero. Asymptotically, the stiffness matrix requires $O(N^2)$ memory to store and each matrix-vector multiplication (and so each Krylov subspace iteration) has $O(N^2)$ computational complexity. Alternatively, a direct solver has $O(N^3)$ computational complexity to invert the linear system at each time step.

This has profound impact computationally. For example, for a three-dimensional analogue of the nonlocal diffusion model (1) or the peridynamic model (3), each time we refine the mesh size and time step by half, the memory requirement would increase 64 times, and the computational complexity by a direct solver increases by $8^3 \times 2 = 1024$ times!

If we use a finite difference or finite element method with 1000 by 1000 by 1000 nodes to solve an integer-order diffusion PDE or elasticity model in three space dimensions, a multigrid solver can solve the problem on the order of 10^9 operations. However, if we use a similar method with the same number of nodes and time steps to solve a three-dimensional analogue of the nonlocal diffusion model (1) or the peridynamic model (3), a direct solver would have a computational complexity on the order 10^{27} per time step. This will take state-of-the-art supercomputer at least years of CPU time to finish the simulation, to say the least. Here we have not taken into account for the extra computational complexity due to the nonlinearity and the evaluation and assembly of the stiffness matrices in the numerical simulation. The significantly increased computational complexity and memory requirements of the nonlocal diffusion model and the peridynamic model significantly limit their applications in reality, especially for problems in three space dimensions.

A simplified peridynamic model was proposed to reduce its computational complexity by assuming that the horizon parameter $\delta = O(h)$ (Chen and Gunzburger 2011). Under this assumption, the stiffness matrix of a numerical scheme becomes local like a high-order finite difference or finite element method for an integer-order PDE. However, it is not clear why physically relevant horizon parameter δ is proportional to the computationally chosen mesh size h . This inconsistency is also reflected in the error estimate of the corresponding finite element method, which now reduces to the following first-order estimate in this case:

$$\|u - u^h\|_{L^2} \leq Ch \|u\|_{H^2}. \quad (6)$$

In this chapter we go over the recent developments of fast and accurate, matrix-free numerical methods with an efficient matrix evaluation and assembly for the nonlocal diffusion model and the peridynamic model without the assumption of $\delta = O(h)$.

Fast Numerical Methods for a One-Dimensional Peridynamic Model

In this section we go over the development of fast and accurate numerical methods with efficient matrix assembly for a one-dimensional steady-state peridynamic model of a finite bar that consists of microelastic materials, which is given by

the following pseudo-differential equation (Chen and Gunzburger 2011; Silling 2000):

$$\int_{\alpha}^{\beta} \frac{u(x) - u(y)}{|x - y|^{1+\gamma}} dy = f(x), \quad x \in (\alpha, \beta), \tag{7}$$

which is closed by the homogeneous Dirichlet boundary condition that is imposed on the complement of the interval (α, β) or simply in the neighborhood of the interval (α, β) .

A Fast Galerkin Finite Element Method on a Uniform Partition

Define a uniform spatial partition on (α, β) by

$$x_i := \alpha + ih, \quad 0 \leq i \leq N, \quad h := \frac{\beta - \alpha}{N} \tag{8}$$

for a given positive integer N . Let S^h be the linear finite element space defined on $[\alpha, \beta]$ with respect to the given partition.

Let $\psi(\xi) := 1 - |\xi|$ for $\xi \in [-1, 1]$ or zero otherwise. Let $\{\phi_i\}_{i=1}^{N-1}$ be the standard nodal basis functions for S^h , i.e.,

$$\phi_i(x) := \psi\left(\frac{x - x_i}{h}\right) \tag{9}$$

Let $u^h \in S^h$ be the finite element approximation to the true solution u of problem (7). Then u^h can be expressed as follows:

$$u^h(x) = \sum_{j=1}^{N-1} u_j \phi_j(x). \tag{10}$$

Then the Galerkin finite element method can be expressed as follows:

$$\sum_{j=1}^{N-1} a(\phi_i, \phi_j) u_j = l(\phi_i), \quad 1 \leq i \leq N - 1, \tag{11}$$

where the bilinear form $a(\cdot, \cdot)$ and the linear functional l are defined by

$$\begin{aligned} a(\phi_j, \phi_i) &= \frac{1}{h} \int_{\alpha}^{\beta} \phi_i(x) \int_{\alpha}^{\beta} \frac{\phi_j(x) - \phi_j(y)}{|x - y|^{1+\gamma}} dy dx, \\ l(\phi_i) &= \frac{1}{h} \int_{\alpha}^{\beta} f(x) \phi_i(x) dx. \end{aligned} \tag{12}$$

Let $u := [u_1, u_2, \dots, u_{N-1}]^T$, $f := [f_1, f_2, \dots, f_{N-1}]^T$, and $A := [a_{i,j}]_{i,j=1}^{N-1}$ with $a_{i,j} = a(\phi_j, \phi_i)$ and $f_i = l(\phi_i)$. The finite element method (11) can be expressed in a matrix form

$$Au = f. \tag{13}$$

It can be proved that the stiffness matrix A is a symmetric and positive definite matrix. Moreover, it is clear from (12) that the stiffness matrix A is full. Consequently, the Galerkin finite element method (11) has a significantly increased memory requirement and computational complexity compared to that for an integer-order PDE model.

Theorem 1. *The linear system (14) can be solved via the conjugate gradient method in a matrix-free fashion, which requires $O(N)$ memory and has an $O(N \log N)$ computational complexity per conjugate gradient iteration.*

Outline of proof. We carefully studied all the entries of the stiffness matrix A in Wang and Wang (2010) and proved that the stiffness matrix A can be decomposed as

$$A = A_{tr} + T. \tag{14}$$

Here A_{tr} is a tridiagonal matrix and $T = [t_{i,j}]_{i,j=1}^{N-1}$ contains all remaining nonzero entries of A and is a symmetric Toeplitz matrix. We let q_{j-i} denote the common entry in the $(j - i)$ th descending diagonal of T from left to right. Namely,

$$t_{i,j} = q_{j-i}, \quad j \geq i. \tag{15}$$

Then the symmetric Toeplitz matrix A_o can be embedded into a $2(N - 1) \times 2(N - 1)$ circulant matrix C as follows (Chan and Ng 1996):

$$C := \begin{bmatrix} T & S \\ S & T \end{bmatrix} \quad S := \begin{bmatrix} 0 & q_{N-2} & \dots & q_2 & q_1 \\ q_{N-2} & 0 & q_{N-2} & \dots & q_2 \\ \vdots & q_{N-2} & 0 & \ddots & \vdots \\ q_2 & \vdots & \ddots & \ddots & q_{N-2} \\ q_1 & q_2 & \dots & q_{N-2} & 0 \end{bmatrix}. \tag{16}$$

The circulant matrix C can be diagonalized by the discrete Fourier transform matrix (Chan and Ng 1996)

$$C = F^{-1} \text{diag}(Fc) F \tag{17}$$

where c is the first column vector of C and F is the $2(N - 1) \times 2(N - 1)$ discrete Fourier transform matrix.

Table 1 The efficiency of Gauss, CG and FCG

h	Gauss	CG		FCG	
	CPU	Iter.	CPU	Iter.	CPU
2^{-10}	47 s	36	1.19 s	36	0.18 s
2^{-14}	6 days 11 h	46	9 m 50 s	46	3.30 s
2^{-16}	Stopped test	53	3 h 9 m	53	18 s
2^{-17}	Out of memory	N/A	Out of memory	56	1 m 29 s
2^{-18}	N/A	N/A	N/A	58	2 m 51 s
2^{-28}	N/A	N/A	N/A	92	3 days 11 h

For any $v \in \mathbb{R}^{N-1}$, define v_{2N-2} by

$$v_{2N-2} = \begin{bmatrix} v \\ 0 \end{bmatrix}, \quad C v_{2N-2} = \begin{bmatrix} T & S \\ S & T \end{bmatrix} \begin{bmatrix} v \\ 0 \end{bmatrix} = \begin{bmatrix} T v \\ S v \end{bmatrix}. \tag{18}$$

By (18) to evaluate $T v$ for any vector $v \in \mathbb{R}^{N-1}$, we need only to evaluate $C v_{2N}$ for v_{2N-2} defined in (7). Note that the latter can be evaluated using the decomposition (17) via the fast Fourier transform (FFT) in $O((2N) \log(2N)) = O(N \log N)$ operations. Hence, $T v$ can be evaluated in $O(N \log N)$ operations. Note that the matrix-vector multiplication is carried out using (17) and (18), which does not require the entire matrix T , but the first column vector c of the matrix C . Hence, it is matrix-free and requires only $O(N)$ memory.

The matrix A_{Tr} is a tridiagonal matrix which can be stored in $O(N)$ memory, and $A_{Tr} v$ can be evaluated in $O(N)$ operations. \square .

The following table shows the numerical performance of Gaussian elimination (Gauss), the conventional conjugate gradient method (FG), and the fast conjugate gradient method (FCG) for a one-dimensional model problem (Wang and Wang 2010). The numerical results show that all the three solvers generate solutions with the same accuracy. The table below shows the efficiency of different solvers (Table 1).

A Fast Collocation Method on a Graded Mesh

Note that peridynamic model intends to describe problems with singularity. Hence, a fast numerical method on a uniform mesh is very efficient but not as effective. In Tian et al. (2013) we developed a fast collocation method for the peridynamic model (7) with $r = 0$ and an assumed singularity point at the right-end point $x = \beta$. In this case, a geometrically graded mesh is defined below

$$x_0 := \alpha, \quad x_i := \alpha + \sum_{k=1}^i \frac{\beta - \alpha}{2^k}, \quad 1 \leq i \leq N - 1, \quad x_N := \beta. \tag{19}$$

Let $\{\phi_i\}_{i=1}^{N-1}$ be the nodal basis functions defined in (9) except that the partition is graded now. In this case, the trial function u_h is still defined in (18). If we choose x_i ($i = 1 \dots, N - 1$) to be the collocation points, then a collocation method can be formulated as follows:

$$\int_{\alpha}^{\beta} \frac{u_i - u(y)}{|x_i - y|} dy = f(x_i), \quad 1 \leq i \leq N - 1. \tag{20}$$

Then the collocation method (20) can be expressed in the matrix form (14) except that now $f_i = f(x_i)$ and

$$a_{i,j} = \int_{\alpha}^{\beta} \frac{-\phi_j}{|x_i - y|} dy, \quad i \neq j, \quad a_{i,i} = \int_{\alpha}^{\beta} \frac{1 - \phi_i}{|x_i - y|} dy, \quad 1 \leq i \leq N - 1. \tag{21}$$

Since the partition (19) is geometrically graded, the basis functions $\{\phi_i\}_{i=1}^{N-1}$ are not translation invariant. We analyzed the structure of the stiffness matrix A and proved the following theorem in

Theorem 2. *The stiffness matrix A defined in (21) can be decomposed as*

$$A = T + L. \tag{22}$$

Here T is a Toeplitz matrix of order $(N - 1)$ which is obtained by extending the matrix $A_{i,l}$ to an $(N - 1)$ -by- $(N - 1)$ matrix, and L is a low-rank matrix of rank at most 2 which is nonzero only in the last row or last column.

Consequently, A can be stored in $O(N)$ memory. The linear system (14) can be solved via any Krylov subspace iterative method with $O(N \log N)$ operations per iteration.

Outline of proof. Note that the rightmost cell has the same size as its left neighbor, so the gridding violates the geometrically decreasing pattern which affects the structure of the matrix. To analyze the matrix structure, we express the stiffness matrix A as a two-by-two block matrix of the form

$$A = \begin{bmatrix} A_{l,l} & A_{l,r} \\ A_{r,l} & A_{r,r} \end{bmatrix}. \tag{23}$$

Here $A_{l,l}$ is an $(N - 2)$ -by- $(N - 2)$ submatrix that consists of the first $N - 2$ rows and columns in A , $A_{l,r}$ consists of the first $(N - 2)$ entries in the last column of A , $A_{r,l}$ consists of the first $(N - 2)$ entries in the last row of A , and $A_{r,r}$ consists of the entry in the last row and column of A .

We proved in Tian et al. (2013) that the submatrix $A_{l,l}$ is a Toeplitz matrix of order $(N - 1)$. By extending $A_{l,l}$ to a Toeplitz matrix T of order $(N - 1)$, we can get the decomposition (22). The rest of the theorem can be proved as in Theorem 1. \square

For the graded mesh (19), the stiffness matrix A is nonsymmetric. Hence, a nonsymmetric Krylov subspace iterative method should be used. Numerical experiments were presented in Tian et al. (2013) which shows the utility of the fast method on graded mesh compared that on the uniform mesh. This coincides with the observations in the context of the numerical methods for fractional PDEs (Jia et al. 2014; Jia and Wang 2015).

A Fast Collocation Method for the Nonlocal Diffusion Model (1)

In this section we outline the development of a fast bilinear collocation method for the nonlocal diffusion model (1) on a rectangular domain $\Omega = (0, x_r) \times (0, y_r)$. As we focus on the study of the matrix structure, we focus on the steady-state analogue of problem (1). Without loss of generality, we assume that $g \equiv 0$ as it does not affect the structure of the stiffness matrix A .

Let the positive integers N_x and N_y denote the numbers of intervals in the x and y directions, respectively. We define a spatial partition on $\bar{\Omega}$ by

$$\begin{aligned} x_i &:= ih_x \text{ for } i = 0, 1, \dots, N_x \text{ with } h_x := \frac{x_r}{N_x}, \\ y_j &:= jh_y \text{ for } j = 0, 1, \dots, N_y \text{ with } h_y := \frac{y_r}{N_y}. \end{aligned} \tag{24}$$

To handle the discretization on the boundary zone Ω_c , we extend the partition to the nodes (x_i, y_j) for $i = -i_x, -i_x + 1, \dots, -1, 0, 1, \dots, N_x, N_x + 1, \dots, N_x + i_x$ and $j = -j_y, -j_y + 1, \dots, -1, 0, 1, \dots, N_y, N_y + 1, \dots, N_y + j_y$. Here

$$i_x := \left\lceil \frac{\delta}{h_x} \right\rceil, \quad j_y := \left\lceil \frac{\delta}{h_y} \right\rceil \tag{25}$$

are the ceilings of δ/h_x and δ/h_y , respectively.

The two-dimensional pyramid functions $\phi_{i,j}(x, y)$ centered at (x_i, y_j) can be expressed as

$$\begin{aligned} \phi_{i,j}(x, y) &:= \psi\left(\frac{x - x_i}{h_x}\right)\psi\left(\frac{y - y_j}{h_y}\right), \quad (x, y) \in (\Omega \cup \Omega_c) \\ 0 &\leq i \leq N_x, 0 \leq j \leq N_y. \end{aligned} \tag{26}$$

The trial function u is of the form

$$u(x, y) = \sum_{i'=1}^{N_x-1} \sum_{j'=1}^{N_y-1} u_{i',j'} \phi_{i',j'}(x, y), \quad (x, y) \in (\Omega \cup \Omega_c). \tag{27}$$

With the prescribed nonlocal homogeneous Dirichlet boundary data, the unknowns are $u_{i,j}$ for $i = 1, 2, \dots, N_x - 1$ and $j = 1, 2, \dots, N_y - 1$. In the numerical scheme we need $N := (N_x - 1)(N_y - 1)$ collocation points. We thus choose the spatial nodes (x_i, y_j) for $i = 1, 2, \dots, N_x - 1$ and $j = 1, 2, \dots, N_y - 1$ as collocation points. This yields the following collocation formulation:

$$\int_{B_\delta(x_i, y_j)} K(x_i - x', y_j - y')(u(x_i, y_j) - u(x', y')) dx' dy' = f(x_i, y_j), \tag{28}$$

$$1 \leq i \leq N_x - 1, \quad 1 \leq j \leq N_y - 1.$$

If we incorporate (27) into (28), we obtain the following collocation scheme:

$$\sum_{i'=1}^{N_x-1} \sum_{j'=1}^{N_y-1} u_{i',j'} \int_{B_\delta(x_i, y_j)} K(x_i - x', y_j - y') (u(x_i, y_j) - \phi_{i',j'}(x', y')) dx' dy' = f(x_i, y_j), \tag{29}$$

$$1 \leq i \leq N_x - 1, \quad 1 \leq j \leq N_y - 1.$$

Let u and f be the N -dimensional vectors defined by

$$u = [u_{1,1}, \dots, u_{N_x-1,1}, u_{1,2}, \dots, u_{N_x-1,2}, \dots, u_{1,N_y-1}, \dots, u_{N_x-1,N_y-1}]^T, \tag{30}$$

$$f = [f_{1,1}, \dots, f_{N_x-1,1}, f_{1,2}, \dots, f_{N_x-1,2}, \dots, f_{1,N_y-1}, \dots, f_{N_x-1,N_y-1}]^T,$$

and $A = [a_{m,n}]_{m,n=1}^N$ be the N -by- N stiffness matrix. Then the collocation method (29) can be expressed in the matrix form (14) with f_m and $A_{m,n}$ being defined by

$$a_{m,n} = \int_{B_\delta(x_i, y_j)} K(x_i - x', y_j - y') (\delta_{m,n} - \phi_{i',j'}(x', y')) dx' dy', \tag{31}$$

$$f_m = f(x_i, y_j).$$

Here $\delta_{m,n} = 1$ for $m = n$ or 0 otherwise. The global indices m and n are related to the indices (i, j) and (i', j') by

$$\begin{aligned}
 m &= (j - 1)(N_x - 1) + i, \quad 1 \leq i \leq N_x - 1, \quad 1 \leq j \leq N_y - 1, \\
 n &= (j' - 1)(N_x - 1) + i', \quad 1 \leq i' \leq N_x - 1, \quad 1 \leq j' \leq N_y - 1.
 \end{aligned}
 \tag{32}$$

It is clear that the stiffness matrix of the collocation method (28) has a block structure. Hence, assembly of the matrix requires the evaluation of $(2i_x + 1)(N_x - 1) - i_x(i_x + 1)$ entries per matrix block and totally $(2j_y + 1)(N_y - 1) - j_y(j_y + 1)$ blocks, which is totally $(2j_x + 1)(2j_y + 1)N - (2j_x + 1)(N_y - 1)j_x(j_x + 1) - (2j_x + 1)(N_x - 1)j_y(j_y + 1) + i_x(i_x + 1)j_y(j_y + 1)$ of entries in the assembly of the stiffness matrix. This is asymptotically $O(N^2)$ of entries as $h_x \rightarrow 0$ and $h_y \rightarrow 0$. Similarly, the collocation method requires asymptotically $O(N^2)$ of memories to store the matrix. In the numerical simulation of nonlocal diffusion models, the assembly of the stiffness matrix often consumes a large portion of CPU time, as the evaluation of each entry can be very costly.

A direct solver has $O(N^3)$ computational complexity asymptotically, while a Krylov subspace iterative method has $O(N^2)$ computational complexity per Krylov subspace iteration. As we have observed from section “Fast Numerical Methods for a One-Dimensional Peridynamic Model,” the key to developing a fast numerical method is to explore the structure of its stiffness matrix.

In Wang and Basu (2012) we analyzed the structure of the stiffness matrix of a finite difference method for a space-fractional PDE in two space dimensions and developed a fast method accordingly. In that analysis we fully utilize the fact that the fractional derivatives in the space-fractional PDE are in the coordinate directions so that the stiffness matrix of the numerical scheme has a tensor product structure.

However, the singular integration operator in the context of the nonlocal diffusion model (1) couples the unknowns in all the directions and so the stiffness matrix of the numerical scheme (29) is not expected to have a tensor product structure. Hence, the analysis is more complicated. We proved the following theorem in Wang et al. (2014).

Theorem 3. *The stiffness matrix A of the numerical scheme (29) has a block-Toeplitz-Toeplitz-block structure with $(N_y - 1)$ -by- $(N_y - 1)$ blocks*

$$A = \begin{bmatrix} T_0 & T_1 & \dots & T_{j_y} & 0 & 0 & \dots & 0 \\ T_{-1} & T_0 & T_1 & \dots & T_{j_y} & 0 & \dots & 0 \\ \vdots & T_{-1} & T_0 & \ddots & \ddots & T_{j_y} & \ddots & \vdots \\ T_{-j_y} & \ddots & \ddots & \ddots & T_1 & \ddots & \ddots & 0 \\ 0 & T_{-j_y} & \ddots & \ddots & T_0 & T_1 & \ddots & T_{j_y} \\ 0 & \ddots & \ddots & \ddots & T_{-1} & T_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & T_1 \\ 0 & 0 & \dots & 0 & T_{-j_y} & \dots & T_{-1} & T_0 \end{bmatrix}. \tag{33}$$

Each block T_j , $-j_y \leq j \leq j_y$, is a Toeplitz matrix of order $N_x - 1$

$$T_j = \begin{bmatrix} t_{0,j} & t_{1,j} & \dots & t_{i_x,j} & 0 & 0 & \dots & 0 \\ t_{-1,j} & t_{0,j} & t_{1,j} & \dots & t_{i_x,j} & 0 & \dots & 0 \\ \vdots & t_{-1,j} & t_{0,j} & \ddots & \ddots & t_{i_x,j} & \ddots & \vdots \\ t_{-i_x,j} & \ddots & \ddots & \ddots & t_{1,j} & \ddots & \ddots & 0 \\ 0 & t_{-i_x,j} & \ddots & \ddots & t_{0,j} & t_{1,j} & \ddots & t_{i_x,j} \\ 0 & \ddots & \ddots & \ddots & t_{-1,j} & t_{0,j} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & t_{1,j} \\ 0 & 0 & \dots & 0 & t_{-i_x,j} & \dots & t_{-1,j} & t_{0,j} \end{bmatrix}. \tag{34}$$

Outline of proof. It is easy to show that if we express the stiffness matrix A as a block matrix of order $(N_y - 1)$, i.e., $A = [B^{j,j'}]_{j,j'=1}^{N_y-1}$, then the matrix A is a block-banded matrix with bandwidth $2j_y + 1$. Similarly, each nonzero matrix block $B^{j,j'} = [b_{i,i'}^{j,j'}]_{i,i'=1}^{N_x-1}$ of order $(N_x - 1)$ is banded with bandwidth $2i_x + 1$.

Using the substitution

$$x' = x_i + x, \quad y' = y_j + y, \tag{35}$$

we can show the relation

$$\phi_{i',j'}(x', y') = \phi_{i'-i,j'-j}(x, y). \tag{36}$$

Accordingly, we can prove that if

$$j'_1 - j_1 = j'_2 - j_2, \tag{37}$$

then we have

$$b_{i,i'}^{j_1,j'_1} = b_{i,i'}^{j_2,j'_2}, \quad 1 \leq i, i' \leq N_x - 1. \tag{38}$$

Namely, $B^{j_1,j'_1} = B^{j_2,j'_2}$. That is, the matrix $A = [B^{j,j'}]_{j,j'=1}^{N_y-1}$ is block-Toeplitz. Similarly, we prove that if

$$i'_3 - i_3 = i'_4 - i_4, \tag{39}$$

then we have

$$b_{i_3,i'_3}^{j,j'} = b_{i_4,i'_4}^{j,j'}. \tag{40}$$

We have thus proved that each block $B^{j,j'}$ is a Toeplitz matrix of bandwidth $2i_x + 1$. □

From Theorem 3, we see that in the fast numerical method, we need only to evaluate and store $(2i_x + 1)(2j_y + 1)$ entries of the matrix. A fast matrix-vector multiplication algorithm can be developed as in section “Fast Numerical Methods for a One-Dimensional Peridynamic Model,” in which each Krylov subspace iteration has computational complexity of $O(N \log N)$. Numerical experiments presented in Wang et al. (2014) showed the utility of the fast method.

A Fast Collocation Method for the Peridynamic Model (3)

As in section “A Fast Collocation Method for the Nonlocal Diffusion Model (1),” we consider the steady-state analogue of problem (3) with the nonlocal homogeneous Dirichlet data. We use the same partition (24) and the basis functions (26). Without loss of generality, we can assume $c = 1$ in (4), as we can always divide both sides of the model by c in this case. Then the x -component $v(x, y)$ and the y -component $w(x, y)$ of the displacement field $\mathbf{u}(x, y)$ can be expressed as

$$\begin{aligned}
 v(x, y) &= \sum_{i'=1}^{N_x-1} \sum_{j'=1}^{N_y-1} v_{i',j'} \phi_{i',j'}(x, y), \\
 w(x, y) &= \sum_{i'=1}^{N_x-1} \sum_{j'=1}^{N_y-1} w_{i',j'} \phi_{i',j'}(x, y).
 \end{aligned}
 \tag{41}$$

We insert (41) into (3) and enforce the resulting equations at the collocation points (x_i, y_j) for $i = 1, 2, \dots, N_x - 1$ and $j = 1, 2, \dots, N_y - 1$ to obtain the following scheme:

$$\begin{aligned}
 &\int_{B_\delta(x_i, y_j)} \frac{(x' - x_i)^2 \left[v_{i,j} - \sum_{i'=1}^{N_x-1} \sum_{j'=1}^{N_y-1} v_{i',j'} \phi_{i',j'}(x', y') \right]}{\left((x' - x_i)^2 + (y' - y_j)^2 \right)^{3/2}} dx' dy' \\
 &+ \int_{B_\delta(x_i, y_j)} \frac{(x' - x_i)(y' - y_j) \left[w_{i,j} - \sum_{i'=1}^{N_x-1} \sum_{j'=1}^{N_y-1} w_{i',j'} \phi_{i',j'}(x', y') \right]}{\left((x' - x_i)^2 + (y' - y_j)^2 \right)^{3/2}} dx' dy' \\
 &= f^v(x_i, y_j),
 \end{aligned}$$

$$\begin{aligned}
& \int_{B_\delta(x_i, y_j)} \frac{(x' - x_i)(y' - y_j) \left[v_{i,j} - \sum_{i'=1}^{N_x-1} \sum_{j'=1}^{N_y-1} v_{i',j'} \phi_{i',j'}(x', y') \right]}{\left((x' - x_i)^2 + (y' - y_j)^2 \right)^{3/2}} dx' dy' \\
& + \int_{B_\delta(x_i, y_j)} \frac{(y' - y_j)^2 \left[w_{i,j} - \sum_{i'=1}^{N_x-1} \sum_{j'=1}^{N_y-1} w_{i',j'} \phi_{i',j'}(x', y') \right]}{\left((x' - x_i)^2 + (y' - y_j)^2 \right)^{3/2}} dx' dy' \\
& = f^w(x_i, y_j), \quad 1 \leq i \leq N_x - 1, \quad 1 \leq j \leq N_y - 1.
\end{aligned} \tag{42}$$

The collocation scheme (42) can be written in the matrix form

$$\mathbf{A}_{2N} \mathbf{u}_{2N} = \mathbf{f}_{2N}. \tag{43}$$

The $2N$ -dimensional unknown displacement vector \mathbf{u}_{2N} and the right-hand side loading vector \mathbf{f}_{2N} of (42) are labeled in their natural order. At each node (x_i, y_j) , the x component $v_{i,j}$ or $f_{i,j}^v$ is put first and the y component $w_{i,j}$ or $f_{i,j}^w$ second:

$$\begin{aligned}
\mathbf{u}_{2N} & := \left[v_{1,1}, w_{1,1}, \dots, v_{N_x-1,1}, w_{N_x-1,1}, v_{1,2}, w_{1,2}, \dots, v_{N_x-1,2}, w_{N_x-1,2}, \right. \\
& \quad \left. \dots, v_{1,N_y-1}, w_{1,N_y-1}, \dots, v_{N_x-1,N_y-1}, w_{N_x-1,N_y-1} \right]^T, \\
\mathbf{f}_{2N} & := \left[f_{1,1}^v, f_{1,1}^w, \dots, f_{N_x-1,1}^v, f_{N_x-1,1}^w, f_{1,2}^v, f_{1,2}^w, \dots, f_{N_x-1,2}^v, f_{N_x-1,2}^w, \right. \\
& \quad \left. \dots, f_{1,N_y-1}^v, f_{1,N_y-1}^w, \dots, f_{N_x-1,N_y-1}^v, f_{N_x-1,N_y-1}^w \right]^T.
\end{aligned} \tag{44}$$

The $2N$ -by- $2N$ stiffness matrix \mathbf{A}_{2N} can be written as the following N -by- N block matrix:

$$\mathbf{A}_{2N} = \begin{bmatrix} \mathbf{A}_{1,1} & \mathbf{A}_{1,2} & \dots & \mathbf{A}_{1,N-1} & \mathbf{A}_{1,N} \\ \mathbf{A}_{2,1} & \mathbf{A}_{2,2} & \dots & \mathbf{A}_{2,N-1} & \mathbf{A}_{2,N} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{A}_{N-1,1} & \mathbf{A}_{N-1,2} & \dots & \mathbf{A}_{N-1,N-1} & \mathbf{A}_{N-1,N} \\ \mathbf{A}_{N,1} & \mathbf{A}_{N,2} & \dots & \mathbf{A}_{N,N-1} & \mathbf{A}_{N,N} \end{bmatrix}, \tag{45}$$

where each entry of \mathbf{A}_{2N} is a two-by-two matrix

$$\mathbf{A}_{m,n} = \begin{bmatrix} a_{m,n}^{v,v} & a_{m,n}^{v,w} \\ a_{m,n}^{w,v} & a_{m,n}^{w,w} \end{bmatrix}, \tag{46}$$

with the entries in $\mathbf{A}_{m,n}$ being defined by

$$\begin{aligned}
 a_{m,n}^{v,v} &:= \int_{B_\delta(x_i,y_j)} \frac{(x' - x_i)^2(\delta_{m,n} - \phi_{i',j'}(x', y'))}{((x' - x_i)^2 + (y' - y_j)^2)^{3/2}} dx' dy', \\
 a_{m,n}^{v,w} &:= \int_{B_\delta(x_i,y_j)} \frac{(x' - x_i)(y' - y_j)(\delta_{m,n} - \phi_{i',j'}(x', y'))}{((x' - x_i)^2 + (y' - y_j)^2)^{3/2}} dx' dy', \\
 a_{m,n}^{w,w} &:= \int_{B_\delta(x_i,y_j)} \frac{(y' - y_j)^2(\delta_{m,n} - \phi_{i',j'}(x', y'))}{((x' - x_i)^2 + (y' - y_j)^2)^{3/2}} dx' dy'.
 \end{aligned} \tag{47}$$

$A_{m,n}^{w,v} = A_{m,n}^{v,w}$. The global indices m and n are related to the indices (i, j) and (i', j') by (32). The entries of the right-hand side vector \mathbf{f}_{2N} are given by

$$f_{i,j}^v = f^v(x_i, y_j), \quad f_{i,j}^w = f^w(x_i, y_j), \quad 1 \leq i \leq N_x - 1, 1 \leq j \leq N_y - 1. \tag{48}$$

We use the similar idea as in section “A Fast Collocation Method for the Nonlocal Diffusion Model (1)” to conduct the analysis. However, due to the coupling between the x component v and the y component w of the displacement \mathbf{u}_{2N} introduced by the kernel in the model, the stiffness matrix \mathbf{A}_{2N} is not block-Toeplitz-Toeplitz-block. To develop a fast method, we analyze the displacement vector \mathbf{u}_{2N} and the stiffness matrix \mathbf{A}_{2N} .

We first decompose the displacement vector \mathbf{u}_{2N} as the following two auxiliary vectors, which replace the even or odd entries of \mathbf{u}_{2N} by 0, respectively:

$$\begin{aligned}
 \mathbf{u}_{2N}^o &:= [v_{1,1}, 0, \dots, v_{N_x-1,1}, 0, v_{1,2}, 0, \dots, v_{N_x-1,2}, 0, \\
 &\quad \dots, v_{1,N_y-1}, 0, \dots, v_{N_x-1,N_y-1}, 0]^T, \\
 \mathbf{u}_{2N}^e &:= [0, w_{1,1}, \dots, 0, w_{N_x-1,1}, 0, w_{1,2}, \dots, 0, w_{N_x-1,2}, \\
 &\quad \dots, 0, w_{1,N_y-1}, \dots, 0, w_{N_x-1,N_y-1}]^T.
 \end{aligned} \tag{49}$$

We remove the zero entries in the above vectors to obtain two N -dimensional vectors \mathbf{v}_N and \mathbf{w}_N . It is clear that

$$\mathbf{u}_{2N} = \mathbf{u}_{2N}^o + \mathbf{u}_{2N}^e. \tag{50}$$

Next, we introduce the N -by- N matrices $\mathbf{A}_N^{v,v}$, $\mathbf{A}_N^{v,w}$, $\mathbf{A}_N^{w,v}$, and $\mathbf{A}_N^{w,w}$ as follows:

$$\mathbf{A}_N^{v,v} := \begin{bmatrix} a_{1,1}^{v,v} & a_{1,2}^{v,v} & \dots & a_{1,N-1}^{v,v} & a_{1,N}^{v,v} \\ a_{2,1}^{v,v} & a_{2,2}^{v,v} & \dots & a_{2,N-1}^{v,v} & a_{2,N}^{v,v} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{N-1,1}^{v,v} & a_{N-1,2}^{v,v} & \dots & a_{N-1,N-1}^{v,v} & a_{N-1,N}^{v,v} \\ a_{N,1}^{v,v} & a_{N,2}^{v,v} & \dots & a_{N,N-1}^{v,v} & a_{N,N}^{v,v} \end{bmatrix}, \tag{51}$$

$$\mathbf{A}_N^{v,w} := \begin{bmatrix} a_{1,1}^{v,w} & a_{1,2}^{v,w} & \cdots & a_{1,N-1}^{v,w} & a_{1,N}^{v,w} \\ a_{2,1}^{v,w} & a_{2,2}^{v,w} & \cdots & a_{2,N-1}^{v,w} & a_{2,N}^{v,w} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{N-1,1}^{v,w} & a_{N-1,2}^{v,w} & \cdots & a_{N-1,N-1}^{v,w} & a_{N-1,N}^{v,w} \\ a_{N,1}^{v,w} & a_{N,2}^{v,w} & \cdots & a_{N,N-1}^{v,w} & a_{N,N}^{v,w} \end{bmatrix}, \quad (52)$$

$$\mathbf{A}_N^{w,w} := \begin{bmatrix} a_{1,1}^{w,w} & a_{1,2}^{w,w} & \cdots & a_{1,N-1}^{w,w} & a_{1,N}^{w,w} \\ a_{2,1}^{w,w} & a_{2,2}^{w,w} & \cdots & a_{2,N-1}^{w,w} & a_{2,N}^{w,w} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{N-1,1}^{w,w} & a_{N-1,2}^{w,w} & \cdots & a_{N-1,N-1}^{w,w} & a_{N-1,N}^{w,w} \\ a_{N,1}^{w,w} & a_{N,2}^{w,w} & \cdots & a_{N,N-1}^{w,w} & a_{N,N}^{w,w} \end{bmatrix}, \quad (53)$$

and $\mathbf{A}_N^{w,v} = \mathbf{A}_N^{v,w}$.

We decompose the matrix-vector multiplication $\mathbf{A}_{2N}\mathbf{u}_{2N}$ as follows:

$$\mathbf{A}_{2N}\mathbf{u}_{2N} = \mathbf{A}_{2N}\mathbf{u}_{2N}^o + \mathbf{A}_{2N}\mathbf{u}_{2N}^e. \quad (54)$$

In addition, if we collect all the odd numbered rows and all the even numbered rows of $\mathbf{A}_{2N}\mathbf{u}_{2N}^o$, we obtain the following matrix-vector multiplications:

$$\mathbf{A}_N^{v,v}\mathbf{v}_N, \quad \mathbf{A}_N^{w,v}\mathbf{v}_N, \quad (55)$$

respectively. Similarly, if we collect all the odd numbered rows and all the even numbered rows of $\mathbf{A}_{2N}\mathbf{u}_{2N}^e$, we obtain the following matrix-vector multiplications:

$$\mathbf{A}_N^{v,w}\mathbf{w}_N, \quad \mathbf{A}_N^{w,w}\mathbf{w}_N, \quad (56)$$

respectively.

In summary, to evaluate the matrix-vector multiplication $\mathbf{A}_{2N}\mathbf{u}_{2N}$ efficiently, we need only to efficiently evaluate the matrix-vector multiplications $\mathbf{A}_N^{v,v}\mathbf{v}_N$, $\mathbf{A}_N^{w,v}\mathbf{v}_N$, $\mathbf{A}_N^{v,w}\mathbf{w}_N$, and $\mathbf{A}_N^{w,w}\mathbf{w}_N$. Similarly, to efficiently store and assemble the stiffness matrix \mathbf{A}_{2N} , we need only to efficiently store and assemble the submatrices $\mathbf{A}_N^{v,v}$, $\mathbf{A}_N^{w,v}$, $\mathbf{A}_N^{v,w}$, and $\mathbf{A}_N^{w,w}$. We use the similar analysis as in section “[A Fast Collocation Method for the Nonlocal Diffusion Model \(1\)](#)” to prove that the matrices $\mathbf{A}_N^{v,v}$, $\mathbf{A}_N^{w,v}$, $\mathbf{A}_N^{v,w}$, and $\mathbf{A}_N^{w,w}$ are block-Toeplitz-Toeplitz-block matrices. So fast matrix-free numerical methods with efficient storage and evaluation of matrix entries can be developed (Zhang and Wang 2016). Numerical experiments show the utility of the method.

Concluding Remarks

In this chapter we outlined the recent developments of fast numerical methods for linear nonlocal diffusion models and bond-based peridynamic models in one and two space dimensions. We show how the analysis was carried out to take the full advantage of the structure of the stiffness matrices of the numerical methods in its storage, evaluation, and assembly and in the efficient solution of the corresponding numerical schemes. This would significantly reduce the computational complexity and storage of the numerical methods over conventional ones, without using any lossy compression. For instance, we would use the same numerical quadratures for conventional methods to evaluate the singular integrals in the stiffness matrices, except that we only need to evaluate $O(N)$ of them instead of $O(N^2)$ of them. Numerical results are presented to show the utility of these fast methods.

Technically, the contributions of this chapter are summarized as follows: (i) the development of the fast numerical methods for the nonlocal diffusion model (1) and peridynamic model (3) is motivated by that of the fast finite difference methods for FPDEs (Chen et al. 2015; Wang and Basu 2012; Wang and Wang 2010; Yang et al. 2011, 2014). However, the nonlocal diffusion model (1) and peridynamic model (3) give rise to the coupling in all the directions, in contrast to FPDEs in which the fractional derivatives are expressed only in the coordinate directions, and so represent a much stronger coupling than FPDEs do. In this chapter we outline the analysis of the stiffness matrices in the development of fast numerical methods with efficient matrix assembly and storage. (ii) The fast numerical methods developed for FPDEs (Wang and Basu 2012; Wang and Wang 2010; Wang and Du 2013) and nonlocal diffusion models (Wang and Tian 2012; Wang et al. 2014; Wang and Wang 2017) are primarily for scalar problems. However, the peridynamic model (3) is a vector equation. Due to the coupling between the x component v and the y component w of the displacement vector \mathbf{u} introduced by the tensor product kernel, the stiffness matrix in the numerical method is not block-Toeplitz-Toeplitz-block as that for nonlocal diffusion models (Wang et al. 2014). Thus, the fast numerical methods developed for FPDEs and nonlocal diffusion models do not apply to the peridynamic model (3). We outlined the analysis of the coupling effect of the peridynamic model (3) on the stiffness matrix of the numerical method and utilize the structure in the development of a fast collocation method.

We end this section by discussing potential extensions: (i) a time-stepping algorithm needs to be used for time-dependent problems, which actually reduces the condition number of the stiffness matrix due to the impact of the time step size. (ii) The extension to three-dimensional problem is conceptually straightforward as the stiffness matrix should retain a similar structure. However, fractional implementation of a three-dimensional peridynamic model requires tremendous effort. (iii) Additional problems include peridynamic model with variable coefficients, state-based peridynamic model, and numerical methods discretized on locally refined composite meshes.

Acknowledgements This work was supported in part by the OSD/ARO MURI under grant W911NF-15-1-0562 and by the National Science Foundation under grants DMS-1620194 and DMS-1216923.

References

- D. Applebaum, *Lévy Processes and Stochastic Calculus* (Cambridge University Press, Cambridge/New York, 2009)
- H.G. Chan, M.K. Ng, Conjugate gradient methods for Toeplitz systems. *SIAM Rev.* **38**, 427–482 (1996)
- S. Chen, F. Liu, X. Jiang, I. Turner, V. Anh, A fast semi-implicit difference method for a nonlinear two-sided space-fractional diffusion equation with variable diffusivity coefficients. *Appl. Math. Comp.* **257**, 591–601 (2015)
- X. Chen, M. Gunzburger, Continuous and discontinuous finite element methods for a peridynamics model of mechanics. *Comput. Methods Appl. Mech. Eng.* **200**, 1237–1250 (2011)
- K. Dayal, K. Bhattacharya, Kinetics of phase transformations in the peridynamic formulation of continuum mechanics. *J. Mech. Phys. Solids* **54**, 1811–1842 (2006)
- D. del-Castillo-Negrete, B.A. Carreras, V.E. Lynch, Fractional diffusion in plasma turbulence. *Phys. Plasmas* **11**, 3854 (2004)
- M. D’Elia, R. Lehoucq, M. Gunzburger, Q. Du, Finite range jump processes and volume-constrained diffusion problems, Sandia National Labs SAND, 2014–2584 (Sandia National Laboratories, Albuquerque/Livermore, 2014)
- Q. Du, M. Gunzburger, R. Lehoucq, K. Zhou, Analysis and approximation of nonlocal diffusion problems with volume constraints. *SIAM Rev.* **54**, 667–696 (2012)
- Q. Du, L. Ju, L. Tian, K. Zhou, A posteriori error analysis of finite element method for linear nonlocal diffusion and peridynamic models. *Math. Comp.* **82**, 1889–1922 (2013)
- E. Emmrich, O. Weckner, The peridynamic equation and its spatial discretisation. *Math. Model. Anal.* **12**, 17–27 (2007)
- W. Gerstle, N. Sau, S. Silling, Peridynamic modeling of concrete structures. *Nucl. Eng. Des.* **237**, 1250–1258 (2007)
- M. Ghajari, L. Iannucci, P. Curtis, A peridynamic material model for the analysis of dynamic crack propagation in orthotropic media. *Comput. Methods Appl. Mech. Eng.* **276**, 431–452 (2014)
- Y.D. Ha, F. Bobaru, Characteristics of dynamic brittle fracture captured with peridynamics. *Eng. Fract. Mech.* **78**, 1156–1168 (2011)
- J. Jia, C. Wang, H. Wang, A fast locally refined method for a space-fractional diffusion equation. # IE0147, ICFDA’14 Catania, 23–25 June 2014. Copyright 2014 IEEE ISBN:978-1-4799-2590-2
- J. Jia, H. Wang, A preconditioned fast finite volume scheme for a fractional differential equation discretized on a locally refined composite mesh. *J. Comput. Phys.* **299**, 842–862 (2015)
- X. Lai, B. Ren, H. Fan, S. Li, C.T. Wu, R.A. Regueiro, L. Liu, Peridynamics simulations of geomaterial fragmentation by impulse loads. *Int. J. Numer. Anal. Meth. Geomech* **39**, 1304–1330 (2015)
- M.M. Meerschaert, A. Sikorskii, *Stochastic Models for Fractional Calculus* (De Gruyter, Berlin, 2011)
- R. Metzler, J. Klafter, The random walk’s guide to anomalous diffusion: a fractional dynamics approach. *Phys. Rep.* **339**, 1–77 (2000)
- B. Øksendal, *Stochastic Differential Equations: An Introduction with Applications* (Springer, Heidelberg, 2010)
- E. Oterkus, E. Madenci, O. Weckner, S.A. Silling, P. Bogert, Combined finite element and peridynamic analyses for predicting failure in a stiffened composite curved panel with a central slot. *Compos. Struct.* **94**, 839–850 (2012)
- M.L. Parks, R.B. Lehoucq, S.J. Plimpton, S.A. Silling, Implementing peridynamics within a molecular dynamics code. *Comp. Phys. Comm.* **179**, 777–783 (2008)

- M.L. Parks, P. Seleson, S.J. Plimpton, S.A. Silling, R.B. Lehoucq, Peridynamics with LAMMPS: a user guide v0.3 beta, SAND report 2011–8523 (Sandia National Laboratories, Albuquerque/Livermore, 2011)
- I. Podlubny, *Fractional Differential Equations* (Academic Press, San Diego, 1999)
- P. Seleson, Improved one-point quadrature algorithms for two-dimensional peridynamic models based on analytical calculations. *Comput. Methods Appl. Mech. Eng.* **282**, 184–217 (2014)
- P. Seleson, Q. Du, M.L. Parks, On the consistency between nearest-neighbor peridynamic discretizations and discretized classical elasticity models. *Comput. Methods Appl. Mech. Eng.* **311**, 698–722 (2016)
- P. Seleson, D. Littlewood, Convergence studies in meshfree peridynamic simulations. *Comput. Math. Appl.* **71**, 2432–2448 (2016)
- P. Seleson, M.L. Parks, On the role of the influence function in the peridynamic theory. *Int. J. Multiscale Comput. Eng.* **9**, 689–706 (2011)
- P. Seleson, M.L. Parks, M. Gunzburger, R.B. Lehoucq, Peridynamics as an upscaling of molecular dynamics. *Multiscale Model. Simul.* **8**, 204–227 (2009)
- S.A. Silling, Reformulation of elasticity theory for discontinuous and long-range forces. *J. Mech. Phys. Solids* **48**, 175–209 (2000)
- S.A. Silling, E. Askari, A meshfree method based on the peridynamic model of solid mechanics. *Comput. Struct.* **83**, 1526–1535 (2005)
- S.A. Silling, M. Epton, O. Wecker, J. Xu, E. Askari, Peridynamic states and constitutive modeling. *J. Elast.* **88**, 151–184 (2007)
- S.A. Silling, O. Weckner, E. Askari, F. Bobaru, Crack nucleation in a peridynamic solid. *Int. J. Fract.* **162**, 219–227 (2010)
- S. Sun, V. Sundararaghavan, A peridynamic implementation of crystal plasticity. *Int. J. Solids Struct.* **51**, 3350–3360 (2014)
- H. Tian, H. Wang, W. Wang, An efficient collocation method for a non-local diffusion model. *Int. J. Numer. Anal. Model.* **10**, 815–825 (2013)
- X. Tian, Q. Du, Analysis and comparison of different approximations to nonlocal diffusion and linear peridynamic equations. *SIAM J. Numer. Anal.* **51**, 3458–3482 (2013)
- C. Wang, H. Wang, A fast collocation method for a variable-coefficient nonlocal diffusion model. *J. Comput. Phys.* **330**, 114–126 (2017)
- H. Wang, T.S. Basu, A fast finite difference method for two-dimensional space-fractional diffusion equations. *SIAM J. Sci. Comput.* **34**, A2444–A2458 (2012)
- H. Wang, N. Du, Fast alternating-direction finite difference methods for three-dimensional space-fractional diffusion equations. *J. Comput. Phys.* **258**, 305–318 (2013)
- H. Wang, H. Tian, A fast Galerkin method with efficient matrix assembly and storage for a peridynamic model. *J. Comput. Phys.* **231**, 7730–7738 (2012)
- H. Wang, H. Tian, A fast and faithful collocation method with efficient matrix assembly for a two-dimensional nonlocal diffusion model. *Comput. Methods Appl. Mech. Eng.* **273**, 19–36 (2014)
- H. Wang, K. Wang, T. Sircar, A direct $O(N \log^2 N)$ finite difference method for fractional diffusion equations. *J. Comput. Phys.* **229**, 8095–8104 (2010)
- Q. Yang, I. Turner, F. Liu, M. Iis, Novel numerical methods for solving the time-space fractional diffusion equation in 2D. *SIAM Sci. Comput.* **33**, 1159–1180 (2011)
- Q. Yang, I. Turner, T. Moroney, F. Liu, A finite volume scheme with preconditioned Lanczos method for two-dimensional space-fractional reaction-diffusion equations. *Appl. Math. Model.* **38**, 3755–3762 (2014)
- X. Zhang, H. Wang, A fast method for a steady-state bond-based peridynamic model. *Comput. Methods Appl. Mech. Eng.* **311**, 280–303 (2016)
- X. Zhang, M. Gunzburger, L. Ju, Nodal-type collocation methods for hypersingular integral equations and nonlocal diffusion problems. *Comput. Methods Appl. Mech. Eng.* **299**, 401–420 (2016)
- K. Zhou, Q. Du, Mathematical and numerical analysis of linear peridynamic models with nonlocal boundary condition. *SIAM J. Numer. Anal.* **48**, 1759–1780 (2010)



Peridynamic Functionally Graded and Porous Materials: Modeling Fracture and Damage

41

Ziguang Chen, Sina Niazi, Guanfeng Zhang, and Florin Bobaru

Contents

Introduction	1354
Peridynamic Formulation for Functionally Graded Materials (FGMs)	1357
Brief Review of Peridynamic Theory for Elastic Brittle Materials	1358
A Fully Homogenized Peridynamic Model for FGMs	1360
An Intermediate-Homogenization Peridynamic Model for Two-Phase Composites	1361
Numerical Studies for Dynamic Crack Propagation in FGMs	1364
Problem Setting	1364
Boundary and Loading Conditions	1367
Convergence Studies for Crack Propagation Path in Dynamic Brittle Fracture of FGMs	1369
Results and Discussion	1373
Quasi-Static Fracture in Brittle Porous Elastic Materials	1380
Conclusions	1383
Appendix A	1383
Numerical Discretization	1383
References	1385

Abstract

In this chapter, we present two peridynamic models for composite materials: a locally homogenized model (FH-PD model, based on results reported in Cheng

Z. Chen (✉)

Department of Mechanics, Huazhong University of Science and Technology, Wuhan, Hubei Sheng, China

Hubei Key Laboratory of Engineering, Structural Analysis and Safety Assessment, Wuhan, China
e-mail: chenziguang@huskers.unl.edu

S. Niazi · G. Zhang · F. Bobaru

Mechanical and Materials Engineering, University of Nebraska–Lincoln, Lincoln, NE, USA
e-mail: sinaniazi85@yahoo.com; zh.guanfeng@gmail.com; fbobaru2@unl.edu

et al. (Compos Struct 133: 529–546, 2015)) and an intermediately homogenized model (IH-PD model). We use these models to simulate fracture in functionally graded materials (FGMs) and in porous elastic materials. We analyze dynamic fracture, by eccentric impact, of a functionally graded plate with monotonically varying volume fraction of reinforcements. We study the influence of material gradients, elastic waves, and of contact time and magnitude of impact loading on the crack growth from a pre-notch in terms of crack path geometry and crack propagation speed. The results from FH-PD and IH-PD models show the same cracking behavior and final crack patterns. The simulations agree very well, through full failure, with experiments. We discuss advantages offered by the peridynamic models in dynamic fracture of FGMs compared with, for example, FEM-based models. The models lead to a better understanding of how cracks propagate in FGMs and of the factors that control crack path and its velocity in these materials. The IH-PD model has important advantages when compared with the FH-PD model when applied to composite materials with phases of disparate mechanical properties. An application to fracture of porous and elastic materials (following Chen et al. (Peridynamic model for damage and fracture in porous materials, 2017)) shows the major effect local heterogeneities have on fracture behavior and the importance of intermediate homogenization as a modeling approach of crack initiation and growth.

Keywords

Peridynamics · Functionally graded materials · Composite materials · Dynamic fracture · Crack propagation · Impact · Brittle fracture · Porous elastic materials

Introduction

Functionally graded materials (FGMs) are a special class of composite materials characterized by spatially varying material properties with the goals of eliminating/reducing stress concentrations, relaxing residual stresses, and enhancing the bonding strength of the composite constituents. Generally, FGMs are multiphase materials having *continuously varying* volume fractions of constituent phases most often along a desired spatial direction. Benefits of using FGMs or FGM coatings in thermomechanical applications are shown in, for example, Kashtalyan and Menshykova (2009), Bobaru (2007), and Thai and Kim (2015). In designing components involving FGMs, it is important to consider imperfections, such as cracks, often preexisting as manufacturing defects or generated by external loads during service. Fracture mechanics of FGMs, especially their dynamic failure, plays a key role in the design of FGMs structures. The fracture behavior of FGMs has been studied especially in the past two decades (see, e.g., Eischen 1987; Jin and Batra 1996; Delale and Erdogan 1983; Wang et al. 2014; Marur and Tippur 2000; Anlas et al. 2000; Itou 2010). Theoretical studies conclude that the leading term in the crack-tip stress field is of the inverse-square-root-singularity for any form of elastic

modulus variation and the elastic modulus has a significant effect on the crack-tip stress field, while the Poisson's ratio has little effect on it (see Delale and Erdogan 1983). Only recently have studies focused on dynamic fracture problems in FGMs. Dynamic analysis of cracks in FGMs under impact loading conditions has appeared in, e.g., Itou (2010), Guo et al. (2004), and Ma et al. (2005), and analysis of crack propagation in FGMs is also presented in Xia and Ma (2007), Kidane et al. (2010), Cheng and Zhong (2007), Cheng et al. (2010), Lee (2009), and Matbuly (2009), for example.

Due to the lack of symmetry in material properties, fracture is inherently mixed-mode when a crack in a FGM is not parallel to the direction of material properties variation, or when loading is asymmetric relative to the crack plane. Jin et al. (2009) experimentally investigated the quasi-static mixed-mode crack propagation in a FGM beam under offset loading. Abanto-Bueno and Lambros (2006) experimentally studied quasi-static mixed-mode crack initiation and growth in FGM through fracture experiments. Shukla and his coauthors (2006) presented a detailed experimental study to understand the dynamic fracture behavior of FGMs. Kirugulige and Tippur (2006) have conducted mixed-mode dynamic fracture experiments on FGM samples made of compositionally graded glass-filled epoxy plates with initial edge cracks along the material gradient. Rousseau and Tippur (2001) presented an experimental study of the crack-tip deformation and fracture parameter histories in compositionally graded glass-filled epoxy under low velocity impact loading.

Most efforts dedicated to simulating static or dynamic fracture behavior of FGM have used the finite element method (FEM). Kim and Paulino (2004) used local remeshing technique to predict the crack path for the mixed-mode quasi-static fracture test in Rousseau and Tippur (2001). Kirugulige and Tippur (2008) have reexamined the mixed-mode dynamic fracture experiment reported in Kirugulige and Tippur (2006) by using the cohesive-zone finite element method. Zhang and Paulino (2005) have conducted mode-I and mixed-mode dynamic fracture simulations in FGMs by cohesive finite elements. Bayesteh and Mohammadi (2013) have used the extended finite element method (XFEM) to analyze the fracture behavior of orthotropic FGMs. The approaches mentioned above used the classical continuum mechanics models to treat fracture. Such approaches are based on partial differential equations and, for describing discontinuities such as cracks, one has to introduce special modeling techniques (such as adaptive remeshing) that define crack surfaces as new boundaries of the updated domain. Continuum-type methods using the cohesive FEM and XFEM require special criteria to decide, for example, when to branch a dynamic crack (see Ha and Bobaru 2011; Song et al. 2008), as well as explicit tracking of the crack path. Brittle damage processes involved in branching of a crack depend on the flow of the strain energy density. A general, predictive theory for how energy is dissipated in the creation of experimentally observed damage around the region where a crack branches (the mirror–mist–hackle transition) is not available. Because of these difficulties, classical models that use a flat surface to represent the actual rough fracture surface are not able to release the correct amount of energy for crack growth in a dynamic setting and,

therefore, are not able to predict experimentally observed crack propagating speeds in dynamic brittle fracture (see Song et al. 2008). Cohesive-zone models need to modify the experimental values of the fracture energy by several factors in order to get propagation velocities in the range of measured ones (see Song et al. 2008).

An alternative approach to modeling dynamic fracture is the new nonlocal continuum model (proposed by Silling (2000) and Silling et al. (2007)), peridynamics, which eliminates spatial derivatives from the formulation with the goal of having a consistent mathematical model for problems with discontinuities in the displacement field. Peridynamics is a reformulation of continuum mechanics in which each material point is connected (via peridynamic bonds) with points in a certain region around it, and not only with its nearest neighbors. In order to overcome mathematical inconsistencies in the classical continuum mechanics models of problems in which cracks initiate and evolve in time, peridynamics uses an integral of forces over a nonlocal region (called horizon) around a point to replace the divergence of the stress tensor in the equations of motion. The model is particularly well suited for dealing with cracks and damage in solid mechanics especially in situations where the crack path is not known in advance.

Peridynamics has been applied to damage and failure analysis of homogeneous and nonhomogeneous materials. Silling and Askari (2005) employed the bond-based peridynamics to study crack growth and impact of a sphere on a brittle target. Bobaru and his coauthors (Ha and Bobaru 2010, 2011; Hu et al. 2012, 2013; Zhang and Bobaru 2015) conducted several peridynamic studies of dynamic fracture in brittle materials and composite materials, including crack branching, and impact loading and fragmentation.

In this chapter, we employ the bond-based peridynamic (PD) model to investigate the dynamic fracture behavior of FGMs under offset (asymmetric) impact loading and understand the factors that influence the crack propagation direction and crack growth in FGMs. We formulate a fully homogenized PD (FH-PD) model for an already homogenized FGM by defining the bond properties between PD nodes based on the elastic and fracture properties at the nodes' locations in the FGM given by experiments. In Cheng et al. (2015), we verified the model in terms of its elastic dynamic behavior (wave propagation induced by sudden loading) by comparing results, in the limit of the nonlocal region going to zero, with analytical ones of the classical mechanics (local) model.

In addition to the FH-PD model, we present a new model in which intermediate homogenization (IH-PD) is used to capture some information about the material microstructure. Different from the FH-PD model, in which the bond properties are computed locally based on an equivalent homogenized composite material, in the IH-PD model we introduce two types of bonds, intraphase bond and interphase bond, that represent the properties of the distinct composite phases and the interfaces between them. For example, for a two-phase composite with phases A and B, we will define A–A bonds (intraphase), B–B bonds (intraphase), and A–B bonds (interphase). The distribution of the bonds depends on the volume fraction of the phases at the nodes forming that particular bond. While the specific geometry of the microstructure is not preserved (unless the scale of the nonlocal region and the

corresponding discretization is at the scale of the smallest geometrical features of the microstructure, rendering the model as an explicit model of the microstructure), the specific volume fraction is, and we shall see how this is critical in modeling failure. The model with this additional information that is preserved from the microscale level will be tested against crack growth in a particular two-phase composite in which one of the phases is void (a poroelastic material).

As examples used to test the models, we analyze dynamic crack growth in FGMs, specifically the mixed-mode dynamic fracture experiment from Kirugullige and Tippur (2006). Both the FH and the IH-PD models are employed to simulate this experiment and we point to similarities and differences between the results. A FGM plate (modeled under 2D plane stress conditions) with monotonically varying volume fraction of reinforcement is simulated under mixed-mode loading by eccentric impact relative to the pre-crack location. The role played by the boundary conditions on the crack path is studied by using two configurations for the impact loading: one with free boundaries and one with loads applied, at the same time as the impact loading, at the support locations, closer, but not identical, to conditions used in experiments. In particular, we notice an interesting phenomenon: duration of impact loading influences the crack path by “attracting” the propagating crack toward the moving location of strain energy concentration, generated by a release surface wave initiated when the load is removed. Both FH-PD and IH-PD models lead to the same cracking patterns for these dynamic loading conditions. One difference between the fully homogenized solution and the intermediately homogenized one is that in the IH-PD model, stress waves become less coherent, being locally dispersed (“noisier” strain energy density maps) by the more detailed representation of the composite microstructure.

While the differences between the two PD models for problems in which extreme dynamics controls the crack behavior (high-intensity impact stress waves) is small, in problems in which the microstructure dictates critical phenomena (like crack initiation), the IH-PD model is able to capture an evolution of the failure not seen with the FH-PD model. Indeed, when a porous material (two-phase composite in which one of the phases is void) is considered under quasi-static loading, the IH-PD model reproduces the experimentally observed dependence of crack initiation and growth on the size of a pre-notch, while the FH-PD model, due to ignoring of any microstructural details, fails to do so. We find that the IH model is as computationally efficient compared with the FH model, which lead us to recommend it against the FH-PD for modeling failure in composite and/or porous materials.

Peridynamic Formulation for Functionally Graded Materials (FGMs)

In this section, we first briefly review the peridynamic theory for elastic brittle materials. Then, we introduce the fully (FH-PD) and intermediate (IH-PH) homogenized peridynamic models for composite materials.

Brief Review of Peridynamic Theory for Elastic Brittle Materials

The peridynamic model is a framework for continuum mechanics based on the idea that pairs of material points exert forces on each other across a finite distance. This concept can be viewed as an effective treatment of material length-scale induced by, for example, the material microstructure. The peridynamic equations of motion for bond-based model are given as (Silling 2000):

$$\rho(\mathbf{x}) \ddot{\mathbf{u}}(\mathbf{x}, t) = \int_{H(\mathbf{x})} f(\mathbf{u}(\mathbf{x}, t) - \mathbf{u}(\hat{\mathbf{x}}, t), \hat{\mathbf{x}} - \mathbf{x}) dV_{\hat{\mathbf{x}}} + \mathbf{b}(\mathbf{x}, t) \tag{1}$$

where f is the pairwise force function in the peridynamic bond that connects point $\hat{\mathbf{x}}$ to \mathbf{x} , \mathbf{u} is the displacement vector field, $\rho(\mathbf{x})$ is the spatially dependent density, and $\mathbf{b}(\mathbf{x}, t)$ is the body force. The integral is defined over a region $H(\mathbf{x})$ called the “horizon region,” or simply the “horizon.” The horizon is the compact supported domain of the pairwise force function around a point \mathbf{x} (see Fig. 1). The horizon region is taken here to be a circle of radius δ . We refer to δ also as the “horizon,” and from the context there should be no confusion whether we refer to the region or its radius.

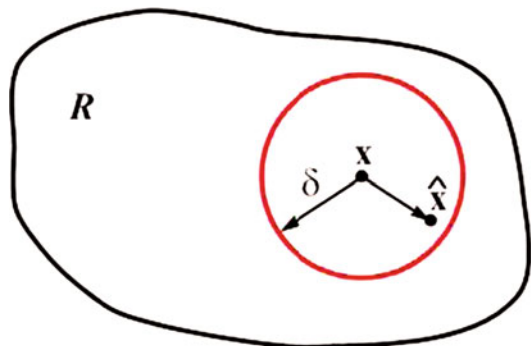
A microelastic material is defined as one for which the pairwise force derives from a potential:

$$f(\boldsymbol{\eta}, \boldsymbol{\xi}) = \frac{\partial \omega(\boldsymbol{\eta}, \boldsymbol{\xi})}{\partial \boldsymbol{\eta}} \tag{2}$$

where $\boldsymbol{\xi} = \hat{\mathbf{x}} - \mathbf{x}$ is the relative position in the reference configuration and $\boldsymbol{\eta} = \hat{\mathbf{u}} - \mathbf{u}$ is the relative displacement of \mathbf{x} and $\hat{\mathbf{x}}$. A micropotential that leads to a linear microelastic material is given by

$$\omega = \frac{1}{2} c(\boldsymbol{\xi}) s^2 \boldsymbol{\xi} \tag{3}$$

Fig. 1 Each point \mathbf{x} interacts directly with the point $\hat{\mathbf{x}}$ in the horizon region delimited by the red circle



where $\xi = \|\xi\|$, $s = \frac{\xi - \xi}{\xi}$ is the relative elongation of a bond, and $\zeta = \|\xi + \eta\|$. The function $c(\xi)$ is called the micro-modulus and has the meaning of bond's elastic stiffness. The integrand in Eq. 1 may have different forms. For examples in diffusion models please see Chen and Bobaru (2015), and in elasticity models please see Chen et al. (2016). It has been observed that the crack propagation speed in brittle fracture is not influenced by the particular shape of the micro-modulus, once the horizon is reasonably small compared to the dimensions of the structure analyzed (see Ha and Bobaru 2010). The pairwise force corresponding to the micropotential given above has the following form:

$$f(\eta, \xi) = \begin{cases} \frac{\xi + \eta}{\|\xi + \eta\|} c(\xi) s & \xi \leq \delta \\ \mathbf{0} & \xi > \delta \end{cases} \tag{4}$$

In this paper, we use the conical 2D plane stress micro-modulus functions (see Ha and Bobaru 2010). Following the same procedure performed to calculate the micro-modulus functions in 1D (see Bobaru et al. 2009), one obtains the conical micro-modulus function in 2D, plane stress conditions (see Ha and Bobaru 2010):

$$c(\xi) = c_1 \left(1 - \frac{\xi}{\delta}\right) = \frac{24E}{\pi\delta^3(1-\nu)} \left(1 - \frac{\xi}{\delta}\right) \tag{5}$$

For a FGM, the elastic modulus depends on the location, in our case on the position of the ends of the peridynamic bond connecting \mathbf{x} and $\hat{\mathbf{x}}$. The model used in this paper uses the location-dependent elastic modulus and this is discussed in the next section.

Failure is introduced in peridynamics by considering that the peridynamic bonds break when they are deformed beyond a critical value, called the critical relative elongation s_0 , computed based on the material's fracture energy. In 2D, the energy per unit fracture length for complete separation of the two halves of the body is the fracture energy G_0 . Equating it to the work done in a PD material to accomplish the separation of the body into two halves gives:

$$G_0 = 2 \int_0^\delta \int_z^\delta \int_0^{\cos^{-1}\left(\frac{z}{\xi}\right)} \left[\frac{c(\xi) s_0^2 \xi}{2} \right] \xi d\theta d\xi dz \tag{6}$$

Substituting the micro-modulus function from Eq. 5 into Eq. 6, s_0 is obtained as:

$$s_0 = \sqrt{\frac{5\pi G_0}{9E\delta}} \tag{7}$$

For the PD-FGM model, G_0 is a function of location of the ends of the peridynamic bond, and this is given in the next section.

A Fully Homogenized Peridynamic Model for FGMs

In this fully homogenized peridynamic (FH-PD) model, a FGM is understood as a material model *locally homogenized* in terms of its elastic properties, density, and fracture energy, corresponding to an actual composition of two phases. The FGM cannot be considered as a globally homogeneous material since its material properties change over large distances, but it can be viewed as a locally homogeneous material because we are not considering the explicit composition of the FGM and we assume that in a sufficiently small neighborhood of any point, material properties are, effectively, constant. We assume that only the volume fraction of each phase influences the elastic and fracture properties of the composite. This assumption is, obviously, a simplification since the particular shape and geometric distribution of the inclusions can affect the elastic and, even more strongly, the fracture properties in the FGM. For problems in which these dependencies are strong, one could first obtain a homogenized peridynamic model from the specific material microstructure. Such a model will lead to a specific horizon size that would capture one or several length-scales generated by the specific microstructure of the composite material (see Silling 2014).

We formulate the peridynamic model for the homogenized FGM concept, and, therefore, the nonlocal region size is not connected to material length-scales generated by a specific microstructure. In this way, we are able to consider the convergence of the nonlocal model solution to the local model, in the limit of the horizon size going to zero (see Silling and Lehoucq 2008; Ha and Bobaru 2010; Bobaru et al. 2009). Such a modeling approach may not be appropriate for modeling failure processes in certain FGMs under certain loading conditions. Part of our goal here is to see to what extent the proposed peridynamic model of a locally homogenized FGM is able to reproduce experimentally observed dynamic fracture behavior for an FGM prepared by varying the volume fraction of small inclusions (glass spheres) in an epoxy matrix as used in Kirugulige and Tippur (2006, 2008). For the type of FGM used in Kirugulige and Tippur (2006, 2008), the fracture behavior under impact conditions reveals a very high repeatability in terms of the crack path (see Fig. 4 in Jain and Shukla (2006)).

Consider two arbitrary material points \mathbf{x} and $\hat{\mathbf{x}}$ connected by a peridynamic bond in an FGM with E_x , ρ_x , G_x and $E_{\hat{x}}$, $\rho_{\hat{x}}$, $G_{\hat{x}}$ being their effective Young's moduli, densities, and fracture energies, respectively. For a peridynamic model of the FGM we need to specify the micro-modulus of the bond connecting these two points. We propose a model in which the material properties used to compute the micro-modulus of the bond connecting points \mathbf{x} and $\hat{\mathbf{x}}$ are the average values of the material properties at the two points. Therefore, in Eqs. 5 and 7 we will use:

$$E(x, \hat{x}) = \frac{(E_x + E_{\hat{x}})}{2}, \quad G_0(x, \hat{x}) = \frac{(G_x + G_{\hat{x}})}{2} \quad (8)$$

The material density is a pointwise quantity and is introduced at each point in the peridynamic equation of motion (see Eq. 1). For the two-dimensional

bond-based peridynamic model of an isotropic and homogeneous material in plane-stress conditions, the Poisson's ratio is $1/3$. For the particular FGM used in the experiments in Kirugulige and Tippur (2006, 2008), Poisson's ratio varies from 0.33 to 0.37, and the authors of Kirugulige and Tippur (2006, 2008) select 0.34 in their modeling and calculations. While Poisson's ratio may influence the path of a growing crack, its variation is not expected to play a significant role in the fracture behavior of FGMs (see Delale and Erdogan (1983) for a discussion on this subject). Peridynamic modeling of other FGMs for which bond-based model's fixed Poisson ratio is not a good fit can be performed using the state-based peridynamic formulation (see Silling and Lehoucq 2008; Silling et al. 2007). The state-based PD formulation eliminates the Poisson ratio restriction, but it comes with a computational penalty, normally increasing the cost by at least a factor of two compared with the bond-based. Because the Poisson ratio for the material considered here is close to $1/3$ and the fact that our primary interest here is to observe the capabilities of a PD model in capturing the evolution of fracture and failure in FGMs and to understand what are the factors that control crack growth in FGMs, in this book chapter, we use the bond-based peridynamic formulation.

An Intermediate-Homogenization Peridynamic Model for Two-Phase Composites

The FH-PD model presented in the previous section assumes a homogenized local material point. In reality, FGMs usually are composites of two or more material phases. The mechanical connections between material points could be interphase or intraphase bonding. An example of a two-phase (A and B) composite is shown in Fig. 2. We consider two arbitrary material points \mathbf{x} and $\hat{\mathbf{x}}$ connected by a peridynamic bond with R and \hat{R} being their Phase A volume fractions, respectively. The peridynamic bond between these two material points has a certain probability

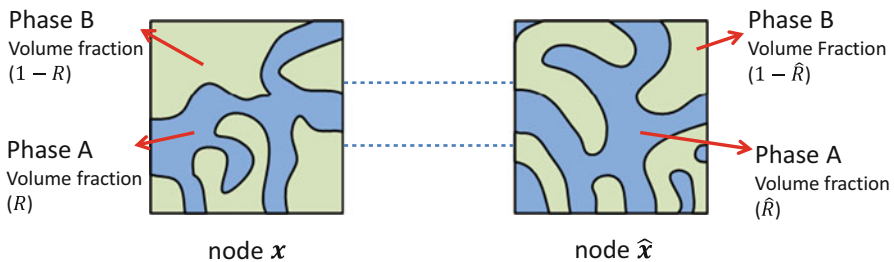


Fig. 2 Description of the intermediate-homogenization PD (IH-PD) model for a two-phase composite material. The bond between the two nodes depends on the local volume fractions of the two composite phases. When the nonlocal size and the discretization size reach the size of the smallest geometrical features of the microstructure, the model matches one in which the microstructural description is explicit

to be considered as a bonding between the same phase or between different phases, depending on the volume fraction of the phases in the two material volumes occupied by the points. We will refer to Phase A – Phase A bonding as A–A bond, Phase B – Phase B bonding as B–B bond, and Phase A – Phase B bonding as A–B bond.

In the new intermediately homogenized peridynamic model (IH-PD, see Chen et al. (2017) for details), we assume a linear relationship between the chance of a bond to be of a certain type and the *local* phase volume fractions at the two nodes. Thus, the bond between material points has chance $R\hat{R}$ to be an A–A bond, chance $(1 - R)(1 - \hat{R})$ to be a B–B bond, and chance $1 - (1 - R)(1 - \hat{R}) - R\hat{R} = R + \hat{R} - 2R\hat{R}$ to be a A–B bond. Heterogeneity is thus introduced by the combination of different mechanical bonds (with different micro-moduli and critical bond strains) connecting at a particular material point.

The detailed steps for the algorithm that implements the IH-PD model, in the model generation stage (preprocessing step), at a node x , are:

- (i) For each of the undetermined bonds in the family of this node, generate a random number r from a uniform distribution in $(0, 1)$.
- (ii) If $r < R\hat{R}$, then label the bond as A–A bond by assigning the micro-modulus and critical bond stretch of material A for this node. If $R\hat{R} < r < (1 - R)(1 - \hat{R}) + R\hat{R}$, then label the bond as B–B bond. Otherwise, label the bond as an A–B bond.

Once the bonds are assigned their type, the mechanical properties (micro-modulus value, critical bond strain) are determined. For A–A bonds, we assign the properties corresponding to phase A, B–B bonds are from phase B, and for A–B bonds we have several options. Two possible ones are based on couplings in series (arithmetic average) or parallel (harmonic average):

$$M_{AB} = \frac{M_A + M_B}{2} \quad (\text{for the a - IH - PD model}),$$

$$\text{and } \frac{2M_A M_B}{M_A + M_B} \quad (\text{for the h - IH - PD model}) \quad (9)$$

where M_A , M_B , and M_{AB} are the mechanical properties corresponding to A, B, and the interphase, phases, respectively. Another possibility for the A–B bonds is to find interface properties from experiments and assign them to A–B bonds. This option is especially useful when the interface is weaker than both of the phases.

Because the algorithm above contains random numbers, every different run that uses a newly seeded random number generator will lead to slightly different results for the same input data. For the elastic model, and a fixed horizon size, however, increasing the m -value (the ratio between the horizon size and the discretization size. See Appendix A for more details of the numerical discretization) results in

attenuating/minimizing these differences. Obviously, with a fixed horizon size, once the node size is smaller than the smallest dominant geometrical feature of the microstructure (pore size in a porous material, fiber diameter in a fiber-reinforced composite, size of debonds or microcracks), the elastic response would no longer change. When the horizon size itself (and the corresponding discretization) is on the scale of the microstructural features, the model essentially becomes an “explicit microstructure” model. Since such models are expensive to compute, the goal for the IH-PD model is to serve as an efficient alternative to the explicit model while incorporating some of the microscale information that has an effect on fracture and failure behavior with the goal of providing more realistic failure response than the FH-PD model could.

To compare the effective properties of the IH-PD model with those from the FH-PD model, we “homogenize” the bonds in the IH-PD model (see Fig. 2) so that each bond has a mechanical property $M(\mathbf{x}, \hat{\mathbf{x}})$ that is computed by considering a combination of parallel bonds with properties M_A , M_B , and M_{AB} between \mathbf{x} and $\hat{\mathbf{x}}$, as follows:

$$M = M_A R \hat{R} + M_B (1 - R) (1 - \hat{R}) + M_{AB} (R + \hat{R} - 2R\hat{R}) \quad (10)$$

By applying Eq. 9, we have:

$$M = \begin{cases} M_B + \frac{1}{2} (M_A - M_B) (R + \hat{R}), & a - \text{IH} - \text{PD} \\ M_B + \frac{(M_A - M_B)[(M_A - M_B)R\hat{R} + M_B(R + \hat{R})]}{(M_A + M_B)}, & h - \text{IH} - \text{PD} \end{cases} \quad (11)$$

Note that for elasticity (and problems in which the microstructure is uniform), the IH-PD and FH-PD converge, in the limit of the horizon going to zero, to the classical composite model solution. However, when the nonlinear model is considered (in which damage is allowed via bond breakage), the results from IH-PD and FH-PD models do not have to be the same. The results from the two models will be different especially for cases with high contrast between the fracture toughness of the composite phases.

When applying the IH-PD model to FGMs with continuously varying volume fraction of reinforcement, taking the horizon size small enough (relative to the overall gradient in composition) leads to $R \cong \hat{R}$. Then, Eq. 11 reduces to:

$$M = \begin{cases} M_B + (M_A - M_B) R(\mathbf{x}), & a - \text{IH} - \text{PD} \\ M_B + \frac{(M_A - M_B)[(M_A - M_B)R^2(\mathbf{x}) + 2M_B R(\mathbf{x})]}{(M_A + M_B)}, & h - \text{IH} - \text{PD} \end{cases} \quad (12)$$

Here, $R(\mathbf{x})$ is the volume fraction of phase A at \mathbf{x} . When $R(\mathbf{x}) = 0$ (100% phase B), $M = M_B$; when $R(\mathbf{x}) = 1$ (100% phase A), $M = M_A$. This is the same as the effective properties in the FH-PD model for an FGM (see Fig. 4 below).

Note that although the effective mechanical properties from the “homogenization” in the IH-PD model could be the same as the ones in FH-PD model, the results

from two models are different and in sections “Numerical Studies for Dynamic Crack Propagation in FGMs” and “Results and Discussion” we will test these differences in both dynamic fracture and quasi-static fracture.

The IH-PD model can also be used for modeling porous elastic materials, by treating the porous material as a composite with one of the phases being a void phase. We apply h-IH-PD model for a porous material. By assuming that phase B is void phase in Fig. 2, material porosity P is the volume fraction of phase B. Substituting $R(x) = 1 - P$ and $M_B = 0$ into the h-IH-PD model in Eq. 12, we obtain the effective mechanical properties as follows:

$$M = M_A(1 - P)^2 \quad (13)$$

Applications of h-IH-PD model for quasi-static fracture in a porous elastic and brittle material are presented in section “Results and Discussion.”

Numerical Studies for Dynamic Crack Propagation in FGMs

Problem Setting

We consider dynamic crack growth in an edge-notch FGM plate specimen under offset impact loading. This setup has been experimentally investigated in Kirugulige and Tippur (2006). The sample geometry and boundary conditions are shown in Fig. 3. We use two types of boundary conditions in our simulations, for two reasons:

- Different dynamic loading induce different wave propagation scenarios that are likely to drive crack growth to different results. These scenarios will allow us to understand the role wave propagation plays in dynamic brittle fracture of FGMs.
- The description given in Kirugulige and Tippur (2006) is not sufficiently detailed to allow us to specify the boundary conditions precisely as those used

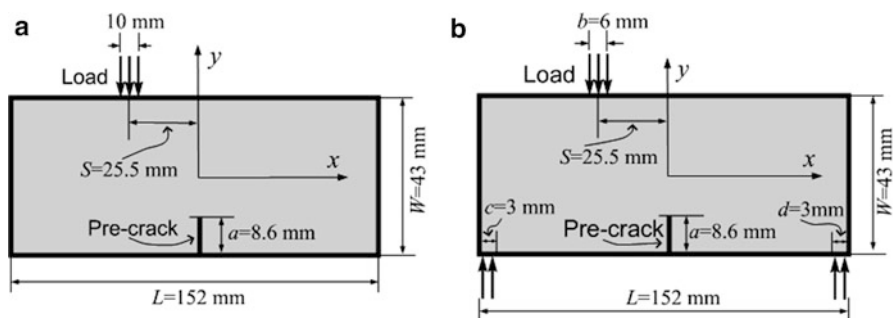


Fig. 3 Specimen geometry and boundary conditions: (a) sudden loading on the top surface, otherwise free boundaries, and (b) sudden loading applied at three points (satisfying static equilibrium)

in the actual experiment. Only the launch speed of the impactor is given. The authors of Kirugulige and Tippur (2006) also mentioned that “the specimen rests on two blocks of soft putty to preclude any interaction from the supports while the specimen undergoes stress wave loading.” The size of the hammer used as the impactor in the experiments is not specified, the rebound time is not given, and the size and properties of the supports are also not reported. The three-point impact loading conditions used in our work, while not identical, are meant to be closer to those in the experiments than the impact loading on the free-free plate.

The cohesive-zone finite element-based model for this problem used in Kirugulige and Tippur (2008) employed only the “free-free” boundary conditions. The computational study in Kirugulige and Tippur (2008) only presents the initial stages of crack propagation, by following the crack growth only for its first 1.5–2.0 cm. The authors of Kirugulige and Tippur (2008) do not explain why the study was limited to this stage of crack growth and why they did not compare the crack paths through final failure of the sample with the experimental results. Perhaps, the FEM solution breaks down after a while. Our peridynamic simulations will track the crack growth through full failure.

The FGM sample used in experiments in Kirugulige and Tippur (2006) is made of epoxy with continuously varying volume fraction of glass filler particles (35 μm mean diameter) from 0 to 40%. The mechanical properties of the constituents of the FGM specimen studied are listed in Table 1. The variation of elastic modulus and mass density along the width of a sample can be found in Fig. 3 in Kirugulige and Tippur (2006). The elastic modulus varies from 10 GPa to 4 GPa over the width and the mass density varies from 1,750 kg/m^3 to 1,175 kg/m^3 over the same width. The FGM specimen is a rectangular plate (see Fig. 3) with width $W = 43$ mm, length $L = 152$ mm, and an edge crack of length $a = 8.6$ mm. The specimen is subjected to an impact loading at an offset distance of $S = 25.5$ mm with respect to the initial crack location.

In their FEM analysis of this problem, the authors of Kirugulige and Tippur (2008) use linear curve fitting to approximate the variation of fracture energy K_{ICR} , elastic modulus E , and the mass density ρ over the sample’s width. We will do the same here, to have a fair comparison between the PD and FEM results. However, in section “Comparison with Experimental and FEM-Based Results: Nonlinear Material Variation” we will also test the crack propagation paths and speed by using the actual, nonlinear variation of elastic modulus and density measured from

Table 1 Nominal bulk properties of the constituent materials used in the FGM samples for the dynamic fracture examples (from Kirugulige and Tippur 2006)

	E (GPa)	ν	ρ (kg/m^3)
Epoxy	3.2	0.34	1,175
Soda-lime glass	70	0.23	2,500

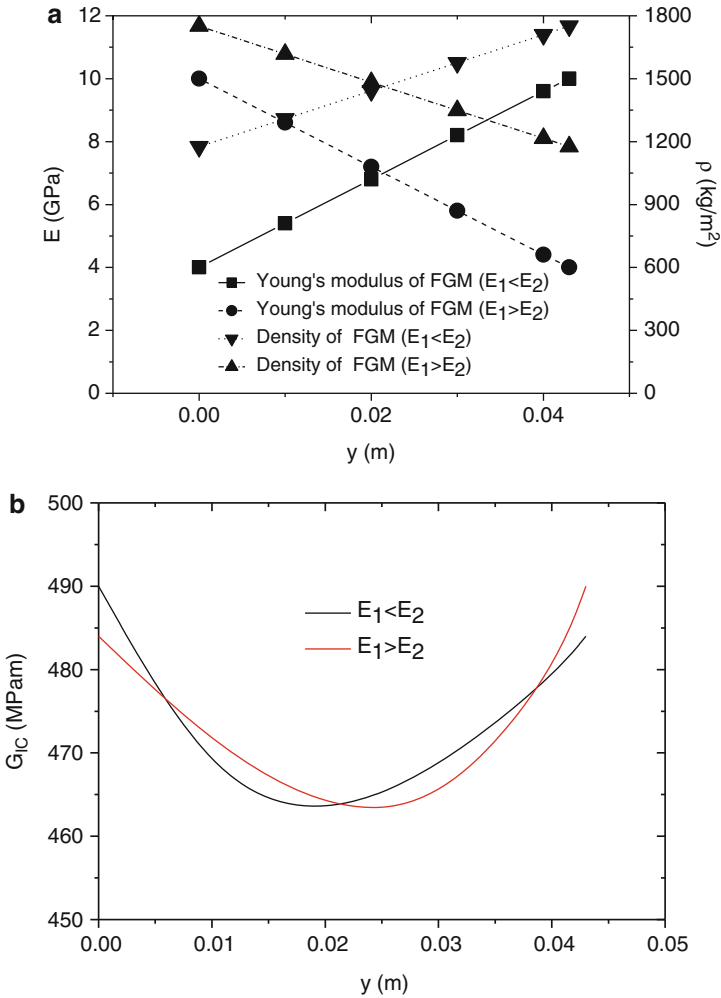


Fig. 4 Linear curve fits for elastic moduli and density across the width of the sample (a). Fracture energy variation across the sample's width based on linear curve fits for the elastic modulus and stress intensity factor (b) (same as those used in the FEM analysis in Kirugulige and Tippur (2008))

experiments and shown in Fig. 3 in Kirugulige and Tippur (2006) and repeated in Fig. 1 in Kirugulige and Tippur (2008).

Let the elastic modulus at the bottom and top surfaces of the specimen be E_1 and E_2 , respectively. The maps with the linear curve fit for density and elastic modulus, as well as the variation of fracture energy along the width of the sample, are shown in Fig. 4.

For the case with the pre-crack on the stiffer side (i.e. $E_1 > E_2$), we have the following linear curve-fits (see Kirugulige and Tippur 2008):

$$K_{ICR}(y) = 2.2 - \frac{2.2 - 1.4}{43}y, \quad 0 \leq y \leq 43 \text{ mm} \quad (14)$$

$$E(y) = 10 - \frac{10 - 4}{43}y, \quad 0 \leq y \leq 43 \text{ mm} \quad (15)$$

$$\rho(y) = 1,750 - \frac{1,750 - 1,175}{43}y, \quad 0 \leq y \leq 43 \text{ mm} \quad (16)$$

$$G_{IC}(y) = \frac{K_{ICR}^2(Y)}{E(y)}, \quad G_{IC}(y) = G_{IIC}(y) \quad (17)$$

Similar linear approximations are used for the other case, when $E_1 < E_2$, when the pre-crack is on the compliant side of the specimen.

The profiles of parameters shown in Eqs. 14, 15, 16, and 17 can be directly applied in Eq. 8 to obtain the bonds information in the peridynamic FH-PD model.

For the IH-PD model, we need the mechanical properties of each phase in the FGMs and their corresponding volume fraction variation over the volume of a peridynamic node, which is decided by the discretization scale used.

In the example here, we consider discretizations that are much larger than the size of the glass beads inclusions. Therefore, we can use the properties as shown above and assume a linear variation of volume fractions. We can thus take the material at the bottom and top surfaces of the sample as being of a single phase. For example, when the pre-crack is on the stiffer side ($E_1 > E_2$), the material at the bottom surface is phase A material, of material parameters shown in Eqs. 14, 15, 16, and 17 with $y = 0$, and the material at the top surface is phase B material, of material parameters shown in Eqs. 14, 15, 16, and 17 with $y = 43$ mm. The volume fraction of each phase is assumed to be linearly distributed in the y -direction. When $E_1 < E_2$, phase A switches with phase B.

Boundary and Loading Conditions

We now present the details about the two types of boundary/loading conditions used in this study. The first type conditions, shown in Fig. 3a, use impact loading of the free beam/plate since this was also used in the finite-element-based computational study in Kirugulige and Tippur (2008). In our case, we apply the normal stress as a uniform distribution of the equivalent body force over a length of 10 mm (from $S - 5$ mm to $S + 5$ mm) on the top surface (see Fig. 3a).

The second type of boundary/loading conditions mimics, to a certain extent, the three-point loading from the experiments. In this case, we apply the dynamic loading, uniformly distributed in space at three locations, corresponding to the two supports and the impact location (see Fig. 3b). The lengths over which the loads are applied are shown in Fig. 3b.

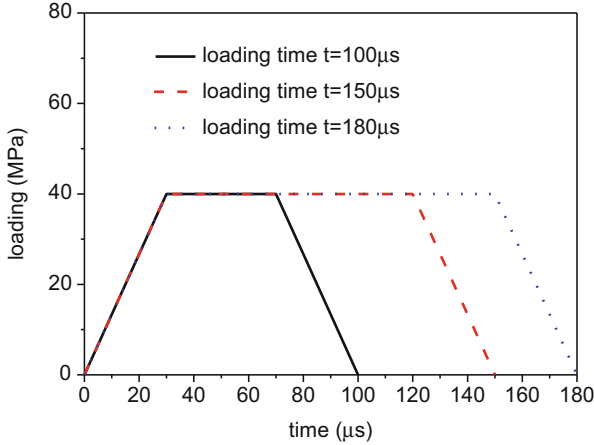


Fig. 5 Variations of the duration of impact loading stress tested in the peridynamic computations

In absence of data from experiments regarding the rebound time of the hammer used to impact the sample, we use a time variation of the impact load described by the following trapezoidal profile (see Fig. 5):

$$\sigma = \begin{cases} \sigma_0 \frac{t}{t_1} & 0 \leq t \leq t_1 \\ \sigma_0 & t_1 \leq t \leq t_2 \\ \sigma_0 \frac{t_3 - t}{t_3 - t_2} & t_2 \leq t \leq t_3 \end{cases} \quad (18)$$

where σ_0 is the maximum impact stress, the ramp-up time $t_1 = 30 \mu\text{s}$, t_3 is the total loading time, and the ramp-down time is $t_3 - t_2 = 30 \mu\text{s}$. We keep the ramp-up and ramp-down times fixed, but we vary the time spent at maximum stress level: $t_2 - t_1 = 40 \mu\text{s}$, $90 \mu\text{s}$, and $120 \mu\text{s}$ resulting in total loading times $t_3 = 100 \mu\text{s}$, $150 \mu\text{s}$, and $180 \mu\text{s}$, respectively (see Fig. 5).

To decide about the actual stress magnitude and corresponding loading force in each of the boundary conditions types, we perform several tests at different loadings and choose the ones that are about the smallest that lead to full failure of the sample (complete separation into two pieces due to crack growth from the pre-notch). For the three-position loading configuration, we select $\sigma_0 = 26 \text{ MPa}$, while for the impact with free boundary conditions, we choose $\sigma_0 = 40 \text{ MPa}$. We also note that some no-fail zones are assigned to the regions around the loading location to avoid immediate failure generation at those locations. In these regions, the critical relative elongation for the peridynamic bonds is set to a large value that is never reached in the present loading conditions. We made sure that these areas do not interfere with the crack growth in any of the computational tests we performed.

For the remaining computational simulations, we use a uniform time step size of 50 ns, which is sufficiently small for all of the tests, based on the stability conditions from Silling and Askari (2005).

Convergence Studies for Crack Propagation Path in Dynamic Brittle Fracture of FGMs

In this subsection, we perform convergence studies in terms of the crack path shape that grows from the pre-cracked FGM specimen under the impact loading for the free–free boundary condition type. Convergence for peridynamic results in *dynamic brittle fracture* has been discussed in Ha and Bobaru (2010, 2011) for isotropic and homogeneous materials and in Hu et al. (2013) for homogeneous but anisotropic materials. Here we study convergence for FGMs with the FH-PD model, and try to answer whether the size of the nonlocal region is sufficiently small, relative to the gradient of material properties function in the FGM, to obtain a crack path that no longer changes when the horizon size decreases further.

Details of the discretized PD model were already given in section “[Numerical Studies for Dynamic Crack Propagation in FGMs](#)” of Cheng et al. (2015) and Appendix A. To those details, we append the combination of Eqs. 7 and 8 for the FH-PD model.

For m -convergence (see Appendix A), we perform tests for a fixed horizon size $\delta = 2$ mm and vary the grid density: $\Delta x = 1$ mm ($m = 2$), 0.5 mm ($m = 4$), and 0.25 mm ($m = 8$). For the case with $E_1 < E_2$ (pre-crack is on the compliant side), we select a total loading time of 150 μs and the resulting crack paths after 210 μs from impact are shown in Fig. 6. The results indicate that, for this horizon size and the particular material variation in the FGM chosen, when a sufficient number of nodes are inside the horizon of the node (when m is larger than 4), no further increases in the grid density changes the crack path. We also note that the lengths of the crack paths in Fig. 6b, c are very similar, meaning that the crack speed is similar between these two cases. The result in Fig. 6a shows that when there is a small number of bonds connecting a node with its nonlocal neighbors ($m = 2$ in 2D produces only 12–20 possible bonds for each node, depending whether nodes covered only partially by the horizon are included or not), the crack path suffers. This is because there are not enough bonds in enough possible directions that can break when stretched beyond their critical value to result in a crack growth that is consistent with the dynamics of the problem. We emphasize here that in peridynamics cracks grow autonomously, because of a cascading of bond rupture events.

For the case when $E_1 > E_2$ (pre-crack is on the stiffer side of the FGM), the m -convergence study is performed using the same numerical grids, the same loading magnitude, and the same total loading time as above. The crack paths for this case at 210 μs after impact are compared in Fig. 7. As in the previous case, it can be observed that the paths with $m = 4$ and 8 are very similar to each other, in both shape and length, meaning that the crack propagation speed between the two grids

Fig. 6 m -convergence in terms of the crack path for an FGM specimen with $E_1 < E_2$ computed with different grids for $\delta = 2$ mm: (a) $m = 2$ ($\Delta x = 1$ mm), (b) $m = 4$ ($\Delta x = 0.5$ mm), and (c) $m = 8$ ($\Delta x = 0.25$ mm). The damage maps shown are at $210 \mu\text{s}$ after impact for the loading case with the total loading time of $150 \mu\text{s}$. Dimensions on the axes are in meters

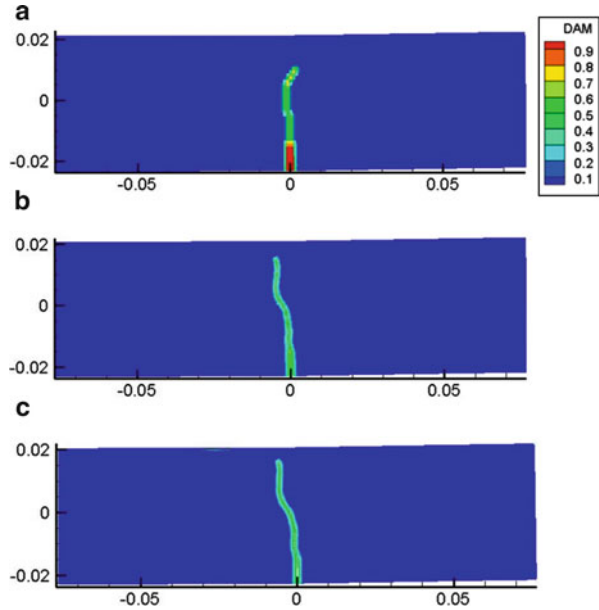
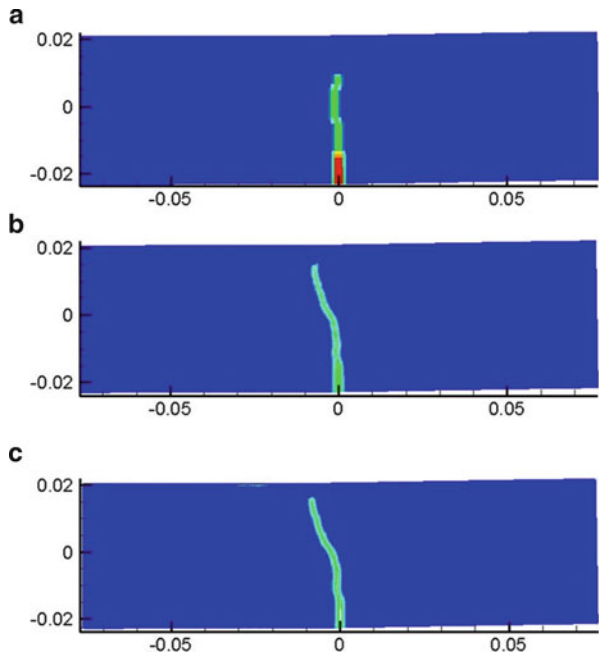
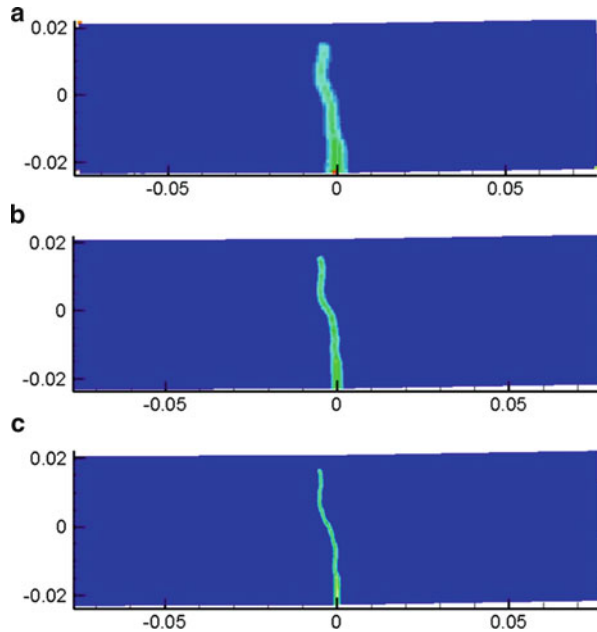


Fig. 7 m -convergence in terms of crack path for an FGM specimen with $E_1 > E_2$ computed with different grids for $\delta = 2$ mm: (a) $m = 2$ ($\Delta x = 1$ mm), (b) $m = 4$ ($\Delta x = 0.5$ mm), and (c) $m = 8$ ($\Delta x = 0.25$ mm). The damage maps shown are at $210 \mu\text{s}$ after impact for the loading case with the total loading time of $150 \mu\text{s}$



are not much different from each other. From these two tests we conclude that, for a given horizon size, a sufficiently high grid density to produce results that no longer change with grid refinement is one produced by $m = 4$. Using a larger value of m

Fig. 8 δ -convergence in terms of the crack path for an FGM specimen with $E_1 < E_2$ computed with different horizon sizes (and grids corresponding to $m = \delta/\Delta x = 4$): (a) $\delta = 4$ mm, (b) $\delta = 2$ mm, and (c) $\delta = 1$ mm. The damage maps shown are at 210 μ s after impact for the loading case with a total loading time of 150 μ s

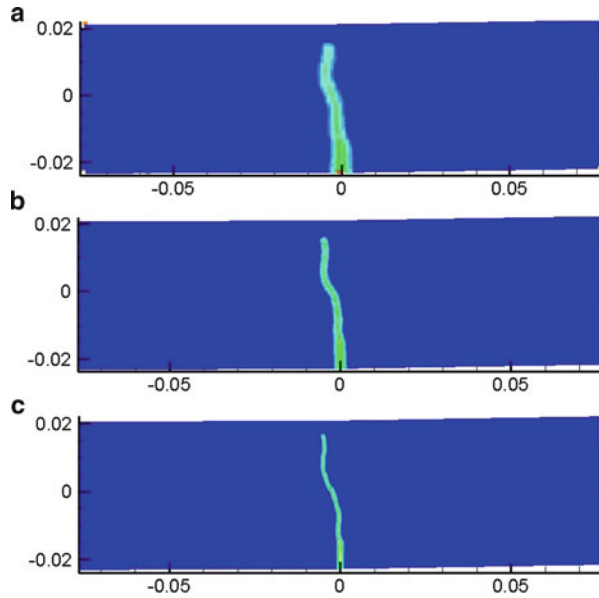


requires higher computational cost, while the results are not affected. Therefore, in the remaining computations we choose $m = 4$. Note also that under m -convergence, the “thickness” of the damage zone does not change (see remarks in Ha and Bobaru (2010)).

To assess whether the crack path is resolved in terms of the horizon size we now turn to δ -convergence (see Appendix A). Since the FH-PD model here is for a locally homogenized FGM model, the horizon size is not connected to an eventual material length-scale generated by specific microstructure architecture of the FGM. Therefore, the role of a δ -convergence study is to determine the largest horizon size that produces a crack path (crack propagation speed) that does not change when a smaller horizon size is used. Since in the examples considered here the cracks are generated from a pre-crack, and because the “size” of the initial damage caused by the pre-crack depends on the size of the horizon (see also Ha and Bobaru (2010) and Bobaru and Hu (2012) for a discussion on this topic), a large horizon size will not represent well the physical dimensions of the pre-crack, leading to stress profiles near the crack tip different from the actual ones. This will likely influence the dynamic crack growth and the crack path in the FGM. Moreover, dynamic fracture is strongly influenced by wave propagation and using a relatively large horizon size leads to a larger nonlocal dispersion for propagating stress waves that influences, in turn, crack growth evolution (Fig. 8).

We perform the δ -convergence study for $m = 4$ and use the following values for the horizon size: 4 mm, 2 mm (as the one used in the m -convergence study above), and 1 mm. Recall that the pre-crack length is close to 9 mm. Since m

Fig. 9 δ -convergence in terms of the crack path for an FGM specimen with $E_1 > E_2$ computed with different horizon sizes (and grids corresponding to $m = \delta/\Delta x = 4$): (a) $\delta = 4$ mm, (b) $\delta = 2$ mm, and (c) $\delta = 1$ mm. The damage maps shown are at 210 μ s after impact for the loading case with a total loading time of 150 μ s



is fixed, each horizon size determines the grid spacing used with it, therefore the corresponding Δx values are: 1 mm, 0.5 mm, and 0.25 mm, respectively. The same impact conditions as those in the m -convergence test are applied. The crack path of both $E_1 < E_2$ and $E_1 > E_2$ at 210 μ s for peridynamic simulations using the three cases specified above are shown in Figs. 9 and 13, respectively. It can be noticed that, as the horizon decreases under δ -convergence, the crack path does not change much, but it becomes more defined, consistent with the assertion that the “spread” or “thickness” of damage is related to the horizon size. Because the crack length at the same time instant is very similar between the last two horizon sizes used, we can also conclude that the crack propagation speed does not change as the horizon changes. Note that the width of the pre-crack mentioned in the experiments in Kirugulige and Tippur (2008, p. 272), is about 0.3 mm, while Fig. 3 in Kirugulige and Tippur (2008) shows the pre-notch width to be in the 1 mm range. Because of this, and because using an even smaller horizon size (while keeping a value of m of at least 4, as required by the results from the m -convergence tests above) would lead to higher computational cost with little difference in the results, in the computations that follow we will use the horizon size $\delta = 1$ mm and a discretization corresponding to $m = 4$, leading to a total of 105,184 nodes for the given geometry.

In the next section, we compare the results from the FH-PD model with those obtained with the IH-PD model for the smallest horizon size used with the FH-PD model.

Results and Discussion

In this section, we first apply the FH-PD model and discuss results in which we vary the time impact loading is applied at its maximum, for the FGM with $E_1 < E_2$ (the results for the case with $E_1 > E_2$ are available in Cheng et al. (2015)). We compare the results with the ones from the IH-PD model and find them to be similar. We also investigate differences in failure response between the free–free conditions and the three-point impact loading conditions. We then discuss the effect of using the actual material variation (nonlinear) versus using the linear curve-fit to those properties when modeling dynamic fracture in FGMs with the FH-PD model. Finally, we compare results with experimental ones on FGM samples and on homogeneous samples.

Influence of Loading Duration and Material Variation: The Crack Path “Attractor”

When the pre-cracked FGM sample with $E_1 < E_2$ is impacted with the 40 MPa equivalent stress in the free–free boundary conditions, the duration over which the sample is under that level of stress loading has a strong influence on the crack path direction growth. A close examination of the strain energy maps during the crack propagation process clearly shows regions where the strain elastic energy density is concentrated and the interaction of stress waves propagating through the structure (see results from FH-PD model, including Movies 1, 2, and 3, for the three loading times of 100 μs , 150 μs , and 180 μs from Fig. 5, respectively). The crack does not start when the first waves reach its tip, but much later when waves reflected from the various boundaries and influenced by the presence of the pre-crack, lead to mixed-mode conditions favorable for crack growth. In Fig. 10 we show some snapshots taken at 144 μs and 210 μs from the initial impact, and after the final failure of the sample, for the three cases of total loading times of 100 μs , 150 μs , and 180 μs (see Fig. 10).

The crack initiates around 120 μs after the impact in all three cases. The initial growth of the crack is close to straight in all cases, but soon after, it starts to deviate from the straight line and it curves to the right when loading time is the shortest (100 μs), while for the other two loading times cases the crack deviates to the left. We explain this phenomenon as follows: for the loading time of 100 μs , the load is removed *before* the crack starts to propagate. The Rayleigh wave generated during the unloading phase, then acts as a “*crack attractor*,” and the crack, “sensing” the strain energy density concentration created by the moving Rayleigh wave on the top surface, starts to follow it and tilts to the right. This explanation is confirmed by the results shown in Fig. 10 and in Movies 1, 2, and 3: for the loading time of 150 μs , the crack tilts left, but when the released Rayleigh wave starts moving to the right, the crack path changes direction and is following the motion of the location with the largest strain energy density on the top surface of the sample, where bending stresses start to become significant due to the opening of the crack. In the case when the loading lasts 180 μs , the crack moves toward the location of the applied load,

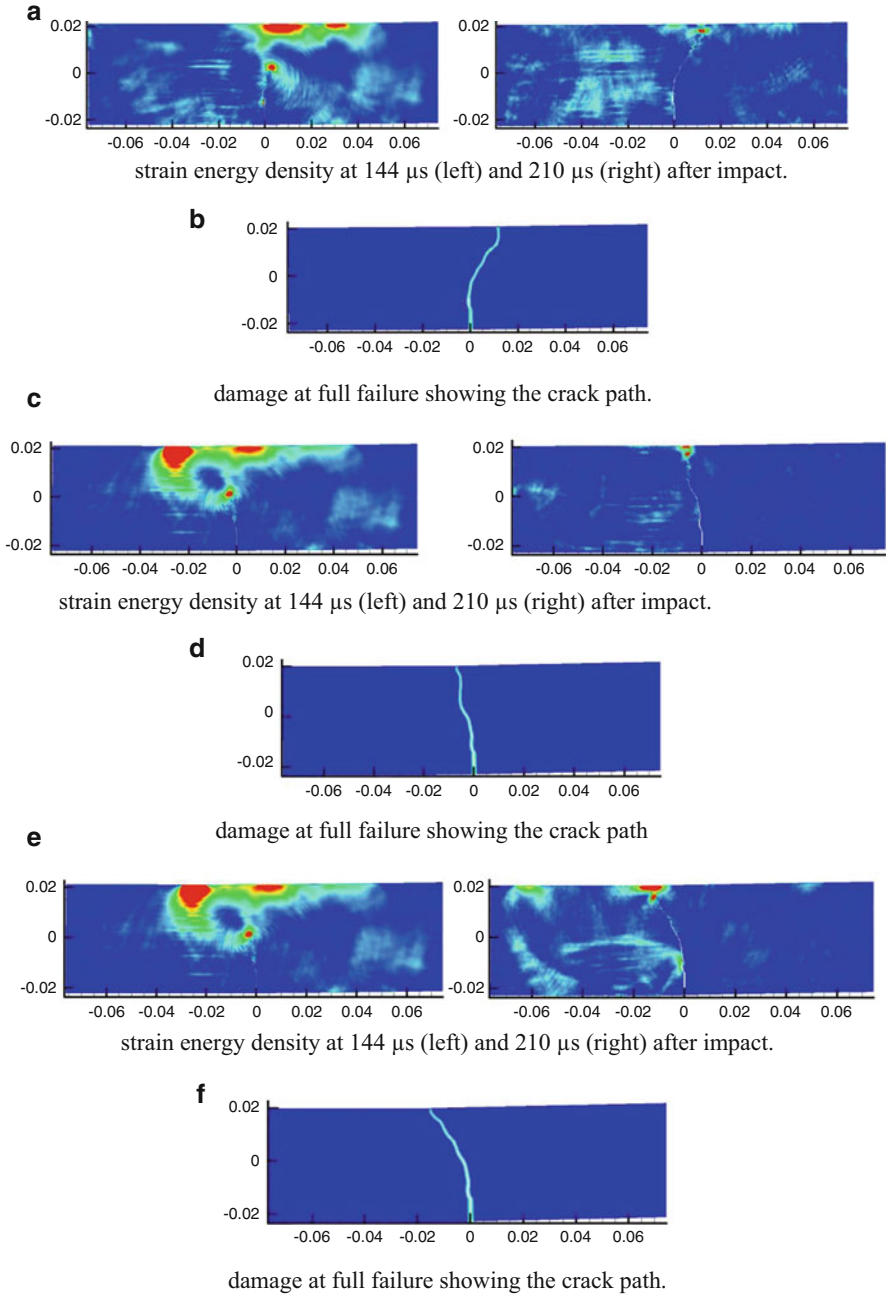


Fig. 10 Snapshots of strain energy density and crack propagation path in the FGM sample with $E_1 < E_2$ at various moments for different total loading times: $t_3 = 100 \mu\text{s}$ for (a) and (b); $t_3 = 150 \mu\text{s}$ for (c) and (d); $t_3 = 180 \mu\text{s}$ for (e) and (f)

and by the time the load is released, the crack path has almost split the sample in two and the Rayleigh wave traveling to the right can change very little of this final stage of failure in the FGM.

The above observations suggest that it might be possible to control the crack path in dynamic fracture from remote locations by using Rayleigh waves.

Using the same loading conditions and solution parameters, the FGM sample with the pre-crack crack on the stiffer side ($E_1 > E_2$) was also tested, and the results are presented in Cheng et al. (2015). The crack initiates now at about 100 μs independent of the time at which the loading stays at the 40 MPa stress value. The explanation presented above for the $E_1 < E_2$ case remains true.

The peridynamic results shown above correctly predict a trend observed in experiments (see Fig. 8 in Kirugulige and Tippur (2006)), namely that the initiation of crack propagation is earlier in the $E_1 > E_2$ case than in the $E_1 < E_2$ case. Notably, this trend *is not* reproduced by the cohesive FEM model in Kirugulige and Tippur (2008). The FEM computations there show either the exact opposite to what is observed in experiments (conversion of stored strain energy into fracture energy starts earlier for the $E_1 < E_2$ case than the $E_1 > E_2$ case, see Sect. 5.2 in Kirugulige and Tippur (2008)), or that the crack initiation time is insensitive to these two different cases of material variation in the FGM (see Sect. 5.4 in Kirugulige and Tippur (2008)).

We also apply IH-PD model, including both a-IH-PD model and h-IH-PD model, for the cases shown in Fig. 10. The comparison of the strain energy density profile between the FH-PD model and the a-IH-PD model at 144 μs is shown in Fig. 11 (see Movies 4, 5, and 6, for strain energy evolution from a-IH-PD model, with the three loading times of 100 μs , 150 μs , and 180 μs , respectively). Only the cases with $t_3 = 100 \mu\text{s}$ and 150 μs are included. The strain energy density profile (at 144 μs) for the case with $t_3 = 180 \mu\text{s}$ is the same as the one with $t_3 = 150 \mu\text{s}$. In both cases, FH-PD model and a-IH-PD model generate the similar strain energy density profiles. One difference between the fully homogenized solution and the intermediately homogenized one, is that in the IH-PD model, stress waves become less coherent, being locally dispersed (“noisier” strain energy density maps) by the more detailed representation of the composite microstructure. Since in this particular impact loading case the damage profile, or the crack propagation, is controlled by the high-intensity impact stress waves, and the two models show similar wave profiles, the final crack patterns from the FH-PD and a-IH-PD models are almost identical, as shown in Fig. 12.

The final damage profiles from a-IH-PD model (see Fig. 12) are the same as the ones from FH-PD model. Only minor differences are seen between the results with the h-IH-PD model and those from the FH-PD model. The reason is that the effective mechanical properties from the a-IH-PD model (see Eq. 11) are the same as the ones in the FH-PD model, while the ones from the h-IH-PD model (see Eq. 11) are slightly smaller than the ones in the FH-PD model. Thus, h-IH-PD model leads to slightly earlier crack initiation and the final crack tip is closer to the right side than the one from FH-PD or a-IH-PD models.

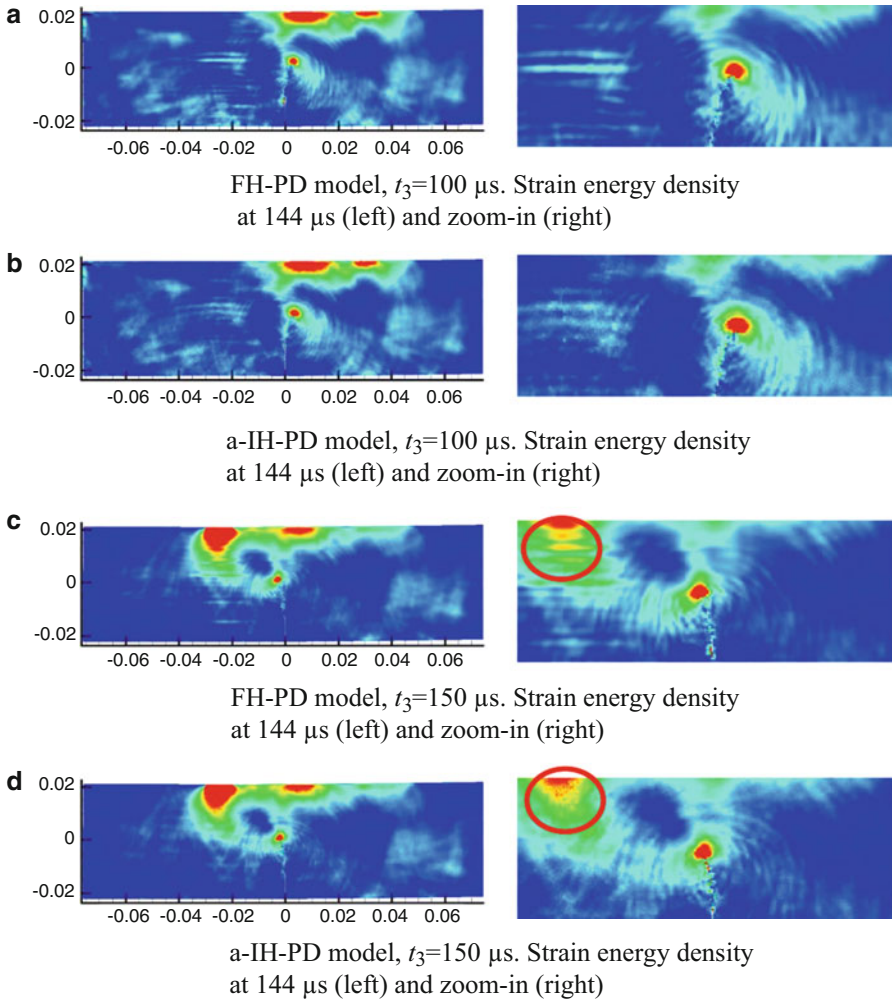


Fig. 11 Snapshots of strain energy density based on the FH-PD model and the a-IH-PD models

Evolution of Crack Propagation Path: Influence of Boundary Conditions

The impact loading used so far is different from the one used in the experiments, where soft supports are holding the impacted structure. In this section, we employ the second type of boundary conditions/loading mentioned in sections “[Problem Setting](#)” and “[Boundary and Loading Conditions](#)” and shown in Fig. 3b. The maximum applied stress (delivered on the peridynamic model as an equivalent body force) in this case needs to be lower than in the previous case in order to get a crack that splits the sample in two pieces because of the bending induced by the three-point impact loading used. We find that using $\sigma_0 = 26 \text{ MPa}$ (see Eq. 17 and Fig. 5) is sufficient to lead to full crack propagation through the width of the sample. The

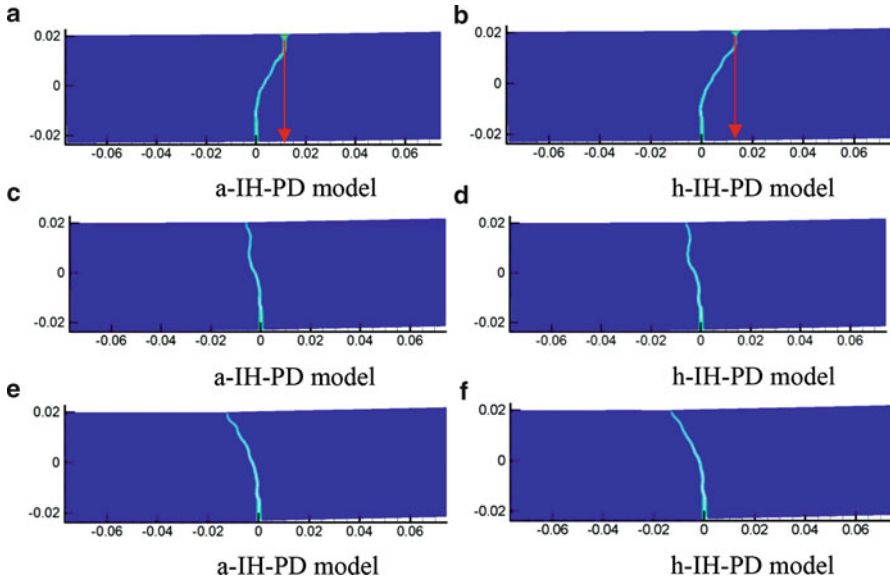


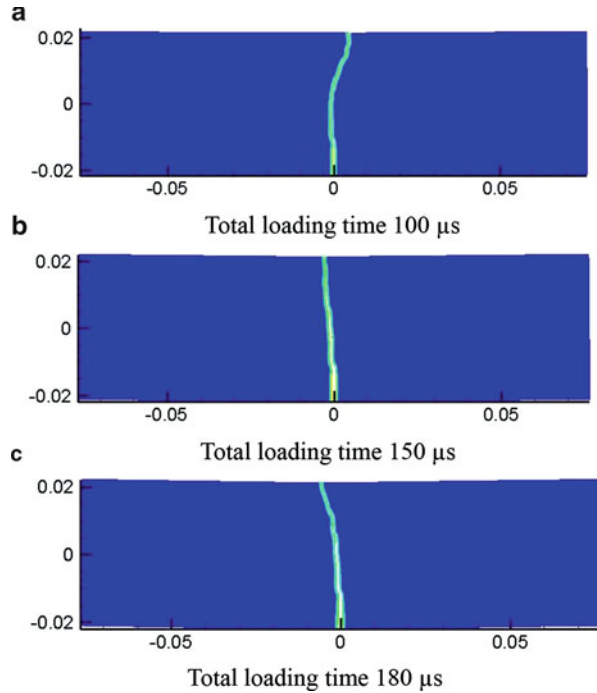
Fig. 12 Damage maps at full failure in the FGM sample with $E_1 < E_2$ for different total loading times: $t_3 = 100 \mu\text{s}$ for (a) and (b); $t_3 = 150 \mu\text{s}$ for (c) and (d); $t_3 = 180 \mu\text{s}$ for (e) and (f)

rest of the computational parameters are identical with those used for the results obtained in the previous section. Note that the impact loads applied on the sides of the sample act simultaneously as the impact load on the top surface. In the actual experiment, the reaction from the sides will be felt only after the waves generated by the striking hammer reach the respective sides. The analysis of the results below will decide whether these differences are significant in terms of dynamic failure in the FGM samples used in our study.

Note that since in the previous section, we have shown that both FH-PD model and IH-PD model lead to similar results, for the remaining of section “[Numerical Studies for Dynamic Crack Propagation in FGMs](#),” we only present results with the FH-PD model.

The crack path for the FGM sample with the pre-crack on the compliant side ($E_1 < E_2$) for total loading times of $100 \mu\text{s}$, $150 \mu\text{s}$, and $180 \mu\text{s}$ is shown in Fig. 13. The initiation time for this case is about $125 \mu\text{s}$, from the moment the loads are applied. The main qualitative difference when comparing simulation results for the three-point loading with free-free impact loading can be seen from the strain energy density movies: once the crack starts moving, the three-point loading leads to the creation of high strain energy density near the center of the top surface induced by the *bending* of the cracked beam. This is the main reason for which the crack path now does not stray much from a straight crack. The reasons for which the crack does not grow completely straight are two: (a) the asymmetric impact loading and (b) the release Rayleigh wave, induced by the unloading phase on the top surface, modifies the location of the “crack attractor” on the top surface.

Fig. 13 Three-point impact loading conditions: damage maps showing the final crack path for the FGM sample with $E_1 < E_2$, under different total loading times. Compare results with those in Fig. 10b, d, and f, respectively



The differences between the impact loading on the free-free beam and the impact at three locations seen above are also seen for the sample with the pre-crack on the stiffer side ($E_1 > E_2$), as shown in Cheng et al. (2015). Interestingly, the differences between the two FGM samples for the total loading times of 150 μs and 180 μs show two trends: when the crack starts in the stiffer region, the crack initiates earlier and the deflection angle (or kink angle) is larger than when the crack starts in the softer region. These results are confirmed by the experimental observations in Kirugulige and Tippur (2006, 2008). A kink angle larger when $E_1 > E_2$ than when $E_1 < E_2$ is also predicted by the maximum tangential stress (MTS) criterion introduced in Erdogan and Sih (1963) for the kink angle in FGMs.

In summary, the three-point impact loading conditions is closer to the actual experimental conditions than the impact on the free-free beam, and the crack paths obtained support this assertion. In addition, our study of varying the total loading time tells us that, most likely, the time that the hammer used in the experiments to impact the structure spends between first touch to rebound is larger than 100 μs . The crack paths obtained for the total loading times of 150 μs and 180 μs are close to those seen from the experiments.

Comparison with Experimental and FEM-Based Results: Nonlinear Material Variation

In this section, we compare the results from the peridynamic simulations with those from the experiments in Kirugulige and Tippur (2006) and from the FEM-based

simulations in Kirugulige and Tippur (2008). The analytical results in Kirugulige and Tippur (2006) and the FEM-based results in Kirugulige and Tippur (2008) used linear curve-fitting of the material properties as shown in Fig. 4. The peridynamic results presented so far used the same curve-fitting. In this section, we directly employ the estimated sigmoidal (described by with a piecewise-linear interpolation) elastic modulus and density measured in Kirugulige and Tippur (2006) and shown in Fig. 3 in that reference. This will determine how strongly the small variations in the elastic modulus and density influence the dynamic crack propagation behavior of FGMs. Finally, we select the peridynamic computational results that most resemble the experimental crack paths to compare among them. We also measure the crack speed growth from our peridynamic solution and compare with it the experimental and FEM-based numerical results from Kirugulige and Tippur (2008).

We extract the data from Fig. 3 in Kirugulige and Tippur (2006) and use it, with the piecewise-linear interpolation shown there, to compute the micro-moduli in our FG-PD model. We apply the three-point impact loading conditions used in the previous section and use the 180 μs total loading time. We obtain the final crack path shapes shown in Fig. 14a for the case with the pre-crack on the compliant side ($E_1 > E_2$), and in Fig. 14b for the case with the pre-crack on the stiffer

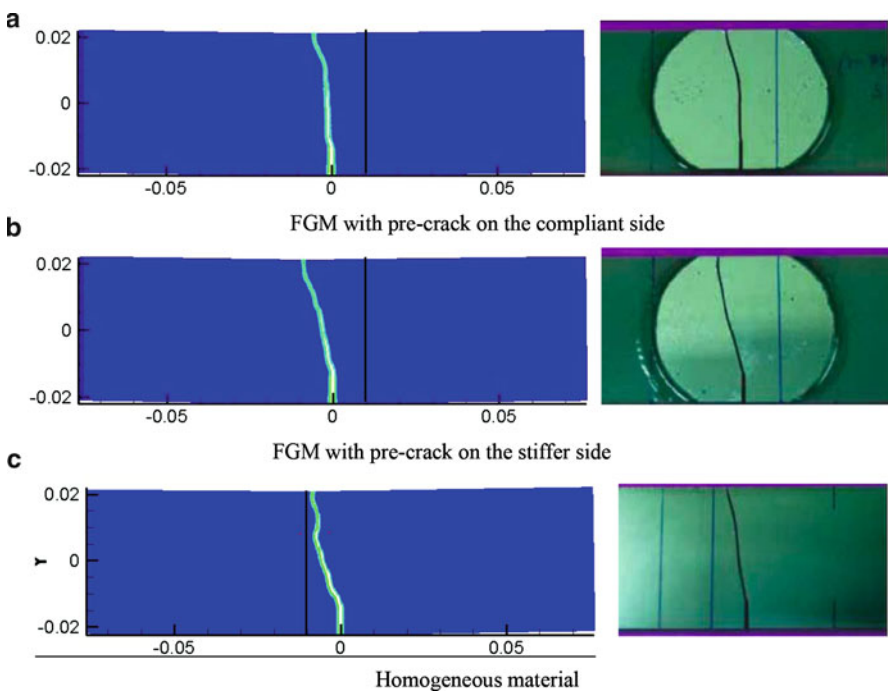


Fig. 14 Damage maps (left column) showing final crack path and experimental results (right column, reproduced from Kirugulige and Tippur (2006), showing only part of the beam). The three-point impact loading was used in the computations with 180 μs total loading time for (a) and (b), and 150 μs for (c). The PD computations for the FGM samples used the nonlinear (experimentally measured in Kirugulige and Tippur (2006)) variation of elastic modulus and density

side of the FGM ($E_1 < E_2$). The crack paths for the corresponding variation of material properties gradation obtained from experiments (taken from Fig. 10 in Kirugulige and Tippur (2006)) are placed next to the computed damage maps. To make it easier to compare the crack shapes, we superpose a black line at 1 cm to the right of the pre-crack tip, similar with what the authors of Kirugulige and Tippur (2006) did. While the dynamic loading used in our simulations is different from that used in the experiments (for the reasons mentioned in the beginning of section “[Quasi-Static Fracture in Brittle Porous Elastic Materials](#)”), we observe a close resemblance between the computed crack paths and the experimental ones. The authors of Kirugulige and Tippur (2006) concluded that: “the differences in the crack paths are attributable directly to the combined effects of elastic gradients as well as fracture toughness gradients.” Our peridynamic model for FGM confirms this statement. Moreover, the peridynamic results shown in section “[Results and Discussion](#)” allow us to add to this statement the following:

- (a) Crack path is influenced by stress wave, including surface waves that can modify the location of the largest bending stresses/strain that “attracts” the crack toward it.
- (b) Stress waves modify the crack path especially after initial crack growth.
- (c) Reflections of elastic waves from the boundaries and their subsequent self-reinforcement determine the time at which the crack starts propagating.

In Fig. 14c, we show the peridynamic result for the homogeneous material with three-point impact loading conditions and a total loading time of 150 μ s. Next to this figure, we show the experimental results for the same material type. We attribute the slight differences in the crack path here to the dynamic loading conditions used in our computations. To make it easier to compare the shape of the crack, we superpose a black line at 1 cm to the left of the pre-crack tip, similar with what the authors of Kirugulige and Tippur (2006) did.

By comparing the results using the nonlinear (piecewise linear) variation in Fig. 14a, b with those using the linear curve-fit for the elastic modulus and density shown in Fig. 13c in this chapter and Fig. 19c in Cheng et al. (2015), we conclude that differences in terms of crack paths exist, but they are small. The main source of these differences are differences in stress waves propagation, caused by the slight differences in local stiffness and density between the two models.

Quasi-Static Fracture in Brittle Porous Elastic Materials

In the previous sections, we have shown that both FH-PD model and IH-PD model can be used to reproduce the dynamic fracture of an FGM plate. While only small differences are observed between the results produced by the FH-PD and IH-PD models when applied to heterogeneous materials in which mechanical properties of the phases are not orders of magnitude different, significant differences are expected when phases differ by orders of magnitude. Examples of such cases are fiber-

reinforced composites, porous materials (in which the voids can be considered as one of the phases), etc. In Chen et al. (2017), we apply both FH-PD model and h-IH-PD model for modeling quasi-static fracture of a porous rock sample. We find that only the IH-PD model delivers similar damage patterns and crack profiles as seen in experiments. In this section, we briefly discuss this test.

The experiments shown in Lin et al. (2009) measure the damage evolution of Berea sandstone notched specimens under quasi-static three-point bending tests. The location of a pre-notch, as a stress concentrator, can control the development of fracture, either through mode I failure for a center-notch beam or mixed-mode failure for an off-center notch beam (Lin et al. 2009). Another critical factor for the development of fracture is the pre-notch length. For the case shown in Fig. 15a (from Lin et al. 2009), under the same loading conditions, different notch lengths lead to different crack patterns. For instance, when the notch is short (5.65 mm), the crack initiates from near the beam center and propagates vertically upward (see section “Results and Discussion” and Fig. 6 in Lin et al. (2009)). When the notch is long (12.23 mm), the beam fails by cracking from the tip of the pre-notch toward the point of loading on the top boundary (see section “Results and Discussion” and Fig. 8 in Lin et al. (2009)).

We apply both the FH-PD and the h-IH-PD models with quasi-static loading conditions to simulate the cases shown in Fig. 15 and to examine the effect the pre-notch length has on the crack pattern. In the FH-PD model of porous material, we use the effective material properties calculated from Eq. 13 for the bond information (micro-modulus and critical bond strain). In the h-IH-PD model, the porous material is treated as a multiphase material with one empty phase (zero modulus). In this model, the bonds involving the empty phase (A–B bond and B–B bond, if phase B is the void phase) bear zero force. These bonds are equivalent with broken bonds. Therefore, the effect of porosity can be represented by inserting “predamage” in the material. The predamage maps corresponding to different porosities, as well as detailed material parameters and model formulation, can be found in Chen et al. (2017).

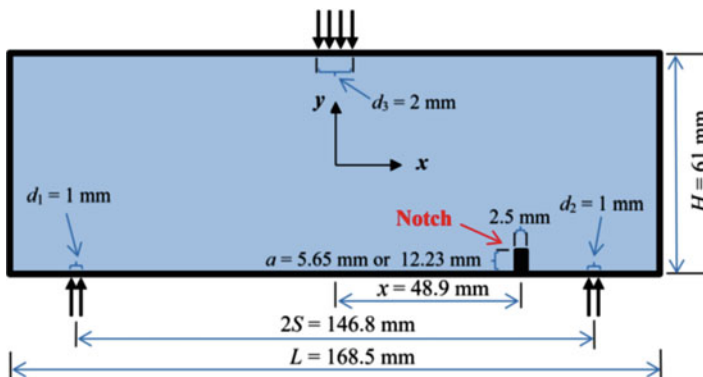


Fig. 15 Geometry and loading configuration used in experiments of Hu et al. (2012) and in the PD simulations below

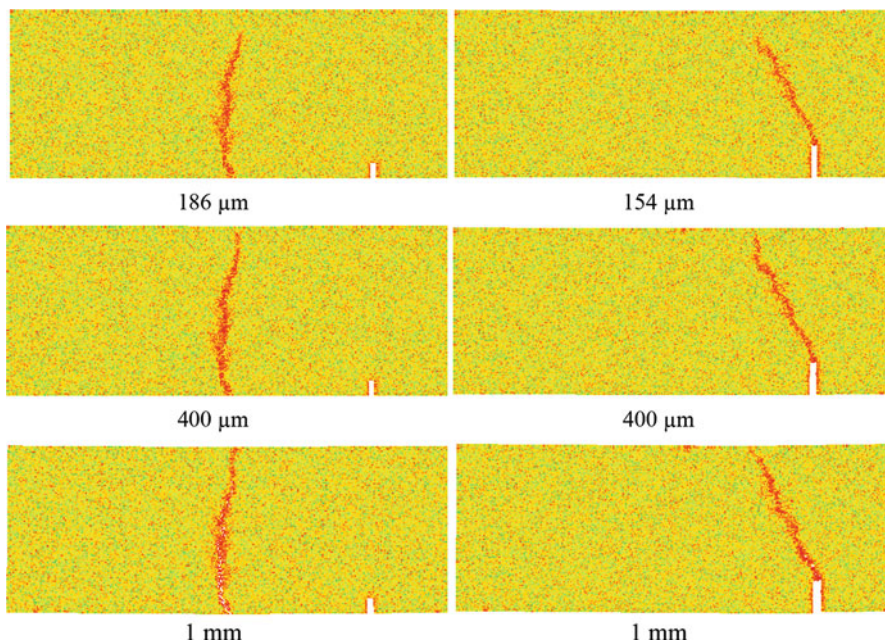


Fig. 16 Evolution of fracture paths (at different imposed displacements) for the short (left) and long (right) pre-notched sample using the h-IH-PD model (see Movies 7 and 8 for the damage evolution for the short and long pre-notched samples). The *color* represents the damage index with the *same color* legend in Fig. 6 (predamage is included in this plot)

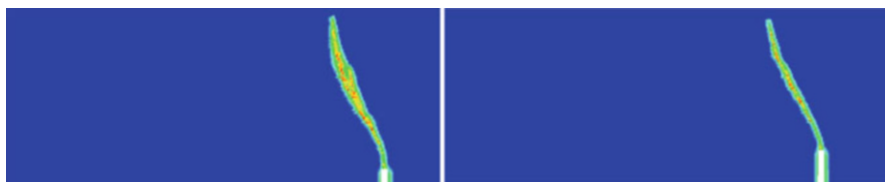


Fig. 17 Fracture paths for the short (left) and long (right) pre-notched samples using FH-PD model (see Movies 9 and 10 for the damage evolution for the short and long pre-notched samples, respectively)

Figure 16 shows the simulation results generated by h-IH-PD model for the developing fracture in the short and long pre-notched Berea sandstone samples. The damage maps show results consistent with the experimental findings on the effect of the notch length: the long pre-notch leads to crack initiating from the pre-notch tip, while in the short pre-notch case, the crack initiates from the bottom-center of the beam.

The FH-PD model does not capture this effect of the pre-notch length on the crack initiation and propagation in the porous sample, as shown in Fig. 17. With the FH-PD model, the crack always initiates from the tip of the pre-notch.

Conclusions

In this chapter, we answered the question of how much homogenization is too much when modeling fracture processes. We presented two peridynamic models, one that homogenizes the material to a greater extent (the fully homogenized peridynamic model, FH-PD) than the other (the intermediately homogenized peridynamic model (IH-PD)). To assess the differences between these models, we studied two types of fracture problems: a dynamic impact problem in a functionally graded material (FGM), for which the crack growth is driven by stress waves, and a quasi-static crack growth problem in a brittle porous elastic material.

The solutions for the problem of mixed-mode dynamic crack propagation in functionally graded glass-filled epoxy matrix obtained with these peridynamic models were similar, and matched closely the experimental results reported in the literature. We provided a detailed computational analysis that showed the influence of loading time and loading conditions on the crack path through full failure of the sample.

In contrast, for a quasi-static crack growth in a porous material, the fully homogenized model fails to capture the experimentally observed fracture behavior in Berean sandstone that is controlled, in this case, by the size of a notch. The intermediately homogenized peridynamic model, however, does reproduce the observed crack growth behavior.

We concluded that for problems in which the microstructure plays an important role on the failure behavior, a fully homogenized strategy will not work. Some information about the microstructure is needed for a predictive model of crack growth in these cases. The new IH-PD model showed that one does not need an explicit description of the small-scale geometry of the material microstructure to predict crack growth in a porous material. Even for the stress-wave driven crack growth in the case of the asymmetric impact loading of the FGM plate, this model produces a landscape of strain energy density that is more realistic compared with the FH-PD model. The crack paths are the same in this case simply because the microstructure effects are rendered secondary by the high flow of energy in the system (stress waves generated by impact and reflecting from the boundaries).

Acknowledgments This work was supported by a grant from ONR (program manager: William Nickerson) and by the AFOSR MURI Center for Material Failure Prediction through Peridynamics (program managers: Dr. David Stargel, Dr. Ali Sayir, Dr. Fariba Fahroo, and James Fillerup). This work was completed utilizing the Holland Computing Center of the University of Nebraska.

Appendix A

Numerical Discretization

The peridynamic equations can be discretized using the finite element method, or any other method appropriate to compute solutions to integro-differential equations

(or integration equations in static model). This approach, however, soon hits well-known obstacles and difficulties for problems with evolving topologies, like those in dynamic fracture and fragmentation. Instead, meshfree-type discretizations are preferred in peridynamics simulations of dynamic failure of materials. The discretization proposed in Silling and Askari (2005) uses the midpoint integration scheme (equivalent to a one-point Gaussian integration) for the domain integral. Numerical simulations are performed using the following discretized equation:

$$\int_{H_x} f(\hat{x} - x, u(\hat{x}, t) - u(x, t)) dV_{\hat{x}} \simeq \sum_{j \in \text{Fam}(i)} c(\xi_{ij}) s_{ij} V_{ij}$$

where $\text{Fam}(i)$ is the family of nodes j with their area (volume in 3D) covered, fully or partially, by the horizon region of nodes i , ξ_{ij} is the bond length between nodes i and j , s_{ij} is the relative elongation for the same bond, and V_{ij} is the area of node j estimated to be covered by the horizon of node i .

Note that node j may not be fully contained within the horizon of node i , so a partial volume integration, first introduced in Hu et al. (2010) and also shown in Zhang and Bobaru (2015), is used here to improve the accuracy of midpoint quadrature scheme. The main advantage of this algorithm compared with one that simply checks whether a node is inside or outside the horizon region is that as the grid density increases (for a fixed horizon value), the numerical convergence (in terms of strain energy density, for example) is monotonic (see Hu et al. 2010).

Both dynamic (see section “Numerical Studies for Dynamic Crack Propagation in FGMs”) and static (see section “Quasi-Static Fracture in Brittle Porous Elastic Materials”) simulations are performed in this work. In the dynamic fracture simulations of the FGM plate (section “Numerical Studies for Dynamic Crack Propagation in FGMs”), we apply Velocity-Verlet method with a time interval of 0.05 μs . For the quasi-static fracture tests in section “Quasi-Static Fracture in Brittle Porous Elastic Materials,” the energy minimization method (see Shewchuk 1994; Zhang et al. 2016; Le and Bobaru 2017) is used, and the conjugate gradient (CG) method with secant line search is adopted to minimize the strain energy of the system. For all static simulations in this chapter, the CG method is used with a convergence tolerance defined by: $\frac{|W_i - W_{i-1}|}{W_{i-1}} < 10^{-6}$, in which W_i and W_{i-1} are the total strain energy at current (i) and previous ($i-1$) CG iterations.

m-convergence and δ -convergence: For a fixed horizon, the ratio $m = \delta/\Delta x$ describes how accurate the numerical quadrature for the integral in Eq. 1 will be. We call this ratio “the horizon factor.” We recall that in the m -convergence we consider the horizon δ fixed and take $m \rightarrow \infty$. The numerical PD approximation will converge to the exact nonlocal PD solution for the given δ . In the case of δ -convergence, the horizon $\delta \rightarrow 0$ while m is fixed or increases with decreasing δ . For δ -convergence and in problems with no singularities, the numerical PD solutions are expected to converge to the classical local solution.

References

- J. Abanto-Bueno, J. Lambros, An experimental study of mixed model crack initiation and growth in functionally graded materials. *Exp. Mech.* **46**, 179–186 (2006)
- G. Anlas, M.H. Santare, J. Lambros, Numerical calculation of stress intensity factors in functionally graded materials. *Int. J. Fract.* **104**, 131–143 (2000)
- H. Bayesteh, S. Mohammadi, XFEM fracture analysis of orthotropic functionally graded materials. *Compos. Part B Eng.* **44**, 8–25 (2013)
- F. Bobaru, Designing optimal volume fractions for functionally graded materials with temperature-dependent material properties. *J. Appl. Mech.* **74**(5), 861–874 (2007)
- F. Bobaru, W. Hu, The meaning, selection, and use of the peridynamic horizon and its relation to crack branching in brittle materials. *Int. J. Fract.* **176**, 215–222 (2012)
- F. Bobaru, M. Yang, L.F. Alves, S.A. Silling, E. Askari, J.F. Xu, Convergence, adaptive refinement, and scaling in 1D peridynamics. *Int. J. Numer. Methods Eng.* **77**, 852–877 (2009)
- Z. Chen, F. Bobaru, Selecting the kernel in a peridynamic formulation: a study for transient heat diffusion. *Comput. Phys. Commun.* **197**, 51–60 (2015)
- Z. Chen, D. Bakenhus, F. Bobaru, A constructive peridynamic kernel for elasticity. *Comput. Methods Appl. Mech. Eng.* **311**, 356–373 (2016)
- Z. Chen, S. Niazi, F. Bobaru, Peridynamic model for damage and fracture in porous materials (2017, in preparation)
- Z.Q. Cheng, Z. Zhong, Analysis of a moving crack in a functionally graded strip between two homogeneous layers. *Int. J. Mech. Sci.* **49**, 1038–1046 (2007)
- Z.Q. Cheng, D.Y. Gao, Z. Zhong, Crack propagating in functionally graded coating with arbitrarily distributed material properties bonded homogeneous substrate. *Acta Mech. Solida Sin.* **23**, 437–446 (2010)
- Z. Cheng, G. Zhang, Y. Wang, F. Bobaru, A peridynamic model for dynamic fracture in functionally graded materials. *Compos. Struct.* **133**, 529–546 (2015)
- F. Delale, F. Erdogan, The crack problem for a nonhomogeneous plane. *J. Appl. Mech.* **50**, 609–614 (1983)
- J.W. Eischen, Fracture of nonhomogeneous material. *Int. J. Fract.* **34**, 3–22 (1987)
- F. Erdogan, G.C. Sih, On the crack extension in plates under plane loading and transverse shear. *Trans. ASME J. Basic Eng.* **84D**(4), 519–525 (1963)
- L.C. Guo, W. LZ, T. Zeng, L. Ma, Fracture analysis of a functionally graded coating-substrate structure with a crack perpendicular to the interface – part II: transient problem. *Int. J. Fract.* **127**, 39–59 (2004)
- Y.D. Ha, F. Bobaru, Studies of dynamic crack propagation and crack branching with peridynamics. *Int. J. Fract.* **162**(1–2), 229–244 (2010)
- Y.D. Ha, F. Bobaru, Characteristics of dynamic brittle fracture captured with peridynamics. *Eng. Fract. Mech.* **78**, 1156–1168 (2011)
- W. Hu, Y.D. Ha, F. Bobaru, Numerical integration in peridynamics, in Technical report, University of Nebraska–Lincoln, Lincoln, 2010
- W. Hu, Y.D. Ha, F. Bobaru, Peridynamic model for dynamic fracture in fiber-reinforced composites. *Comput. Methods Appl. Mech. Eng.* **217–220**, 247–261 (2012)
- W. Hu, Y. Wang, J. Yu, C.F. Yen, F. Bobaru, Impact damage on a thin glass with a thin polycarbonate backing. *Int. J. Impact Eng.* **62**, 152–165 (2013)
- S. Itou, Dynamic stress intensity factors for two parallel interface cracks between a nonhomogeneous bonding layer and two dissimilar elastic half-planes subject to an impact load. *Int. J. Solids Struct.* **47**, 2155–2163 (2010)
- N. Jain, A. Shukla, Mixed mode dynamic fracture in particulate reinforced functionally graded materials. *Exp. Mech.* **46**(2), 137–154 (2006)
- Z.H. Jin, R.C. Batra, Some basic fracture mechanics concepts in functionally graded materials. *J. Mech. Phys. Solids* **44**, 1221–1235 (1996)

- X. Jin, W. LZ, L.C. Guo, Y. HJ, Y.G. Sun, Experimental investigation of the mixed-mode crack propagation in $ZrO_2/NiCr$ functionally graded materials. *Eng. Fract. Mech.* **76**, 1800–1810 (2009)
- M. Kashtalyan, M. Menshykova, Effect of a functionally graded interlayer on three-dimensional elastic deformation of coated plates subjected to transverse loading. *Compos. Struct.* **89**(2), 167–176 (2009)
- A. Kidane, V.B. Chalivendra, A. Shulka, R. Chona, Mixed-mode dynamic crack propagation in graded materials under thermo-mechanical loading. *Eng. Fract. Mech.* **77**, 2864–2880 (2010)
- J.H. Kim, G.H. Paulino, Simulation of crack propagation in functionally graded materials under mixed-mode and non-proportional loading. *Int. J. Mech. Mater. Des.* **1**, 63–94 (2004)
- M.S. Kirugulige, H.V. Tippur, Mixed-mode dynamic crack growth in functionally graded glass-filled epoxy. *Exp. Mech.* **46**(2), 269–281 (2006)
- M.S. Kirugulige, H.V. Tippur, Mixed-mode dynamic crack growth in a functionally graded particulate composite: experimental measurement and finite element simulations. *J. Appl. Mech.* **75**(5), 051102 (2008)
- Q. Le, F. Bobaru, Surface corrections for peridynamic models in elasticity and fracture. *Comput. Mech.* (2017). <https://doi.org/10.1007/s00466-017-1469-1>
- K.H. Lee, Analysis of a transiently propagating crack in functionally graded materials under mode I and II. *Int. J. Eng. Sci.* **47**, 852–865 (2009)
- Q. Lin, A. Fakhimi, M. Haggerty, J.F. Labuz, Initiation of tensile and mixed-mode fracture in sandstone. *Int. J. Rock Mech. Min. Sci.* **46**, 489–497 (2009)
- L. Ma, L.Z. Wu, L.C. Guo, Z.G. Zhou, On the moving Griffith crack in a non-homogeneous orthotropic medium. *Eur. J. Mech. A. Solids* **24**, 393–405 (2005)
- P.R. Marur, H.V. Tippur, Numerical analysis of crack-tip fields in functionally graded materials with a crack normal to the elastic gradient. *Int. J. Solids Struct.* **37**, 5353–5370 (2000)
- M.S. Matbuly, Multiple crack propagation along the interface of a nonhomogeneous composite subjected to anti-plane shear loading. *Meccanica* **44**, 547–554 (2009)
- C.E. Rousseau, H.V. Tippur, Dynamic fracture of compositionally graded materials with cracks along the elastic gradient: experiments and analysis. *Mech. Mater.* **37**, 403–421 (2001)
- J.R. Shewchuk, An introduction to the conjugate gradient method without the agonizing pain, in Technical report, School of Computer Science, Carnegie Mellon University, Pittsburgh, 1994
- S.A. Silling, Reformulation of elasticity theory for discontinuities and long-range forces. *J. Mech. Phys. Solids* **48**, 175–209 (2000)
- S.A. Silling, Origin and effects of nonlocality in a composite. *J. Mech. Mater. Struct.* **9**, 245–258 (2014)
- S.A. Silling, E. Askari, A meshfree method based on the peridynamic model of solid mechanics. *Comput. Struct.* **83**, 1526–1535 (2005)
- S.A. Silling, R.B. Lehoucq, Convergence of peridynamics to classical elasticity theory. *J. Elast.* **93**, 13–37 (2008)
- S.A. Silling, M. Epton, O. Weckner, J. Xu, E. Askari, Peridynamic states and constitutive modeling. *J. Elast.* **88**(2), 151–184 (2007)
- J. Song, H. Wang, T. Belytschko, A comparative study on finite element methods for dynamic fracture. *Comput. Mech.* **42**, 239–250 (2008)
- H.T. Thai, S.E. Kim, A review of theories for the modeling and analysis of functionally graded plates and shells. *Compos. Struct.* **128**, 70–86 (2015)
- Z.H. Wang, L.C. Guo, L. Zhang, A general modeling method for functionally graded materials with an arbitrarily oriented crack. *Philos. Mag.* **94**, 764–791 (2014)
- C.H. Xia, L. Ma, Dynamic behavior of a finite crack in functionally graded materials subjected to plane incident time-harmonic stress wave. *Compos. Struct.* **77**(1), 10–17 (2007)

-
- G. Zhang, F. Bobaru, Why do cracks branch? A peridynamic investigation of dynamic brittle fracture. *Int. J. Fract.* **196**, 59–98 (2015)
- Z. Zhang, G.H. Paulino, Cohesive zone modeling of dynamic failure in homogeneous and functionally graded materials. *Int. J. Plast.* **21**, 1195–1254 (2005)
- G. Zhang, Q. Le, A. Loghin, A. Subramaniyan, F. Bobaru, Validation of a peridynamic model for fatigue cracking. *Eng. Fract. Mech.* **162**, 76–94 (2016)



Numerical Tools for Improved Convergence of Meshfree Peridynamic Discretizations

42

Pablo Seleson and David J. Littlewood

Contents

Introduction	1390
Peridynamic Models and Their Meshfree Discretization	1391
Linearized LPS Model	1393
Discretized Peridynamic Models	1394
Numerical Tools for Improved Convergence	1395
Partial-Volume Algorithms	1396
Influence Functions	1397
Convergence Studies of Static Peridynamic Problems	1398
Preliminary Considerations	1399
Numerical Results	1403
Convergence Studies of Dynamic Peridynamic Problems	1405
Preliminary Considerations	1406
Numerical Results	1408
Concluding Remarks	1413
References	1413

Abstract

Peridynamic models have been employed to simulate a broad range of engineering applications concerning material failure and damage, with the majority of these simulations using a meshfree discretization. This chapter reviews

P. Seleson (✉)

Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA

e-mail: selesonpd@ornl.gov

D. J. Littlewood

Center for Computing Research, Sandia National Laboratories, Albuquerque, NM, USA

e-mail: djlittl@sandia.gov

© This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply 2019

G. Z. Voyiadjis (ed.), *Handbook of Nonlocal Continuum Mechanics for Materials and Structures*, https://doi.org/10.1007/978-3-319-58729-5_39

1389

that meshfree discretization, related issues present in peridynamic convergence studies, and possible remedies proposed in the literature. In particular, we discuss two numerical tools, partial-volume algorithms and influence functions, to improve the convergence behavior of numerical solutions in peridynamics. Numerical studies in this chapter involve static and dynamic simulations for linear elastic state-based peridynamic problems.

Keywords

Peridynamics · Meshfree discretization · Partial volumes · Influence functions · Convergence · Statics · Dynamics

Introduction

Peridynamic models have been employed over the past years to simulate a broad range of engineering applications. These applications include failure and damage in composite laminates (Hu et al. 2012; Kilic et al. 2009; Oterkus and Madenci 2012; Oterkus et al. 2012; Xu et al. 2008), crack propagation and branching in glass (Ha and Bobaru 2010, 2011; Kilic and Madenci 2009), crack nucleation (Littlewood 2011; Silling et al. 2010), impact damage (Bobaru et al. 2012; Seleson and Parks 2011; Silling and Askari 2005; Tupek et al. 2013), fracture in polycrystalline materials (Askari et al. 2008; De Meo et al. 2016; Ghajari et al. 2014), structural health monitoring (Littlewood et al. 2012), and damage in concrete (Gerstle et al. 2007), among many others. There are two main distinctive discretization methods employed by the peridynamic community. These are the finite element method (see, e.g., Chen and Gunzburger 2011) and the meshfree method of Silling and Askari (2005). The former discretization method, in spite of robust supportive mathematical analysis, is not amenable in terms of computational cost and software implementation, particularly for simulations involving complex geometry in three dimensions. Two specific challenges of finite element discretizations of peridynamic models are proper quadratures for peridynamic weak formulations, which double the number of spatial dimensions to be discretized relative to strong formulations, and the necessity to adapt finite element meshes to conform to evolving crack surfaces. In contrast, the latter discretization method, i.e., the meshfree method, has been rapidly adopted by the engineering community, due to its implementation simplicity and its ability to handle material separation in a relatively simple manner. A downside of the meshfree approach is its poor convergence behavior. Specifically, it has been shown in Seleson and Littlewood (2016) that standard meshfree discretizations of peridynamic models exhibit irregular (non-smooth) convergence behavior and a low (average) first-order convergence. Nevertheless, numerical tools have been proposed and studied in Seleson (2014) and Seleson and Littlewood (2016) to correct such irregular convergence behavior, resulting in smoother convergence curves. The (average) first-order convergence is however tied to an underlying piecewise-constant representation of the displacement field in meshfree discretizations and could possibly be improved with higher-order

approximations (as in finite element methods), although these are beyond the scope of this chapter.

Generally speaking, setting up computational studies of convergence of numerical solutions in peridynamics may or may not be a simple task. First of all, as discussed in Bobaru et al. (2009), the concept of convergence in peridynamics has more than one meaning. One may study, for instance, the convergence of a numerical solution of a peridynamic problem to the analytical solution of that problem; in this case, the horizon δ is kept fixed and the grid spacing h is taken to zero. On the other hand, one could study the convergence of a numerical solution of a peridynamic problem to the analytical solution of a corresponding classical (local) problem; in this case, both δ and h are simultaneously taken to zero, most probably at different rates. The various convergence avenues in peridynamics have been illustrated in Du and Tian (2014).

This chapter is dedicated to reviewing the meshfree discretization of peridynamic models, related issues present in convergence studies, and possible remedies proposed in the literature. This chapter also discusses important considerations for setting up convergence studies in peridynamics, such as the choice of horizon, the error norm, the convergence avenue, and the refinement path, as well as related computational costs. We focus on numerical convergence studies for linear elastic state-based peridynamic problems. Following Seleson and Littlewood (2016), we first study the convergence of numerical solutions of static peridynamic problems. Then, we present studies concerning wave propagation in a peridynamic medium to demonstrate how such convergence issues manifest in dynamic problems. To keep this chapter succinct while providing the most impactful results toward practical applications, we restrict the discussion to a select number of three-dimensional problems.

Peridynamic Models and Their Meshfree Discretization

This chapter is concerned with convergence studies in three-dimensional problems for ordinary state-based peridynamic (PD) models, given by the *linear peridynamic solid* (LPS) model (Silling et al. 2007). The LPS strain energy density is

$$W = \frac{1}{2} \left(K - \frac{5G}{3} \right) \vartheta^2 + \frac{1}{2} \frac{15G}{m} (\underline{\omega e}) \bullet \underline{e} \quad (1)$$

and the corresponding force vector state field is given by

$$\underline{\mathbf{T}} = \left(\frac{3K - 5G}{m} \vartheta \underline{\omega x} + \frac{15G}{m} \underline{\omega e} \right) \underline{\mathbf{M}}, \quad (2)$$

where K is the bulk modulus and G is the shear modulus. In Eqs. 1 and 2, ϑ is the dilatation, \underline{x} is the reference position state, \underline{e} is the extension state, m is the weighted volume, $\underline{\mathbf{M}}$ is the deformed direction vector state, and $\underline{\omega} = \omega(|\underline{\xi}|)$ is a

spherical influence function depending on the length of a bond ξ (Seleson and Parks 2011; Silling et al. 2007). For computational purposes, it is useful to express Eqs. 1 and 2, respectively, in a more explicit form as

$$W(\mathbf{x}, t) = \frac{1}{2} \left(K - \frac{5G}{3} \right) (\vartheta[\mathbf{x}, t])^2 + \frac{1}{2} \frac{15G}{m} \int_{\mathcal{H}_x} \omega(|\xi|) (|\mathbf{y}(\mathbf{x} + \xi, t) - \mathbf{y}(\mathbf{x}, t)| - |\xi|)^2 dV_\xi$$

and

$$\begin{aligned} \mathbf{T}[\mathbf{x}, t] \langle \xi \rangle &= \frac{\omega(|\xi|) |\xi|}{m} \left[(3K - 5G) \vartheta[\mathbf{x}, t] + 15G \frac{|\mathbf{y}(\mathbf{x} + \xi, t) - \mathbf{y}(\mathbf{x}, t)| - |\xi|}{|\xi|} \right] \\ &\times \frac{\mathbf{y}(\mathbf{x} + \xi, t) - \mathbf{y}(\mathbf{x}, t)}{|\mathbf{y}(\mathbf{x} + \xi, t) - \mathbf{y}(\mathbf{x}, t)|}, \end{aligned}$$

where

$$\vartheta[\mathbf{x}, t] = \frac{3}{m} \int_{\mathcal{H}_x} \omega(|\xi|) |\xi| (|\mathbf{y}(\mathbf{x} + \xi, t) - \mathbf{y}(\mathbf{x}, t)| - |\xi|) dV_\xi \quad (3)$$

and

$$m = \int_{\mathcal{H}_x} \omega(|\xi|) |\xi|^2 dV_\xi \quad (4)$$

with \mathcal{H}_x the family of \mathbf{x} , and where $\mathbf{y}(\mathbf{x}, t)$ is the position of the material point \mathbf{x} in the deformed configuration at time $t > 0$. For simplicity and for consistency with Seleson and Littlewood (2016), in this chapter the value of m is computed for a point in the bulk of a body and equally assigned to all points in that body, including those near the boundary of the body. This does not affect the static simulations in section “Convergence Studies of Static Peridynamic Problems,” because in those studies m is only required for points within the bulk of the body, but may slightly influence the dynamic results in section “Convergence Studies of Dynamic Peridynamic Problems.” Given a PD body \mathcal{B} , the LPS equation of motion for $\mathbf{x} \in \mathcal{B}$ at time $t \geq 0$ is

$$\begin{aligned} \rho(\mathbf{x}) \frac{\partial^2 \mathbf{u}}{\partial t^2}(\mathbf{x}, t) &= \int_{\mathcal{H}_x} \frac{\omega(|\xi|) |\xi|}{m} \left\{ (3K - 5G) (\vartheta[\mathbf{x}, t] + \vartheta[\mathbf{x} + \xi, t]) \right. \\ &\quad \left. + 30G \frac{|\mathbf{y}(\mathbf{x} + \xi, t) - \mathbf{y}(\mathbf{x}, t)| - |\xi|}{|\xi|} \right\} \\ &\times \frac{\mathbf{y}(\mathbf{x} + \xi, t) - \mathbf{y}(\mathbf{x}, t)}{|\mathbf{y}(\mathbf{x} + \xi, t) - \mathbf{y}(\mathbf{x}, t)|} dV_\xi + \mathbf{b}(\mathbf{x}, t), \end{aligned} \quad (5)$$

where $\boldsymbol{\xi} = \mathbf{x}' - \mathbf{x}$, ρ is the mass density field, \mathbf{u} is the displacement field, and \mathbf{b} is the body force density field.

Linearized LPS Model

For a small deformation, the LPS model can be linearized giving the following strain energy density (Silling 2010)

$$W(\mathbf{x}, t) = \frac{1}{2} \left(K - \frac{5G}{3} \right) (\vartheta^{\text{lin}}[\mathbf{x}, t])^2 + \frac{1}{2} \frac{15G}{m} \int_{\mathcal{H}_x} \frac{\omega(|\boldsymbol{\xi}|)}{|\boldsymbol{\xi}|^2} (\boldsymbol{\xi} \cdot (\mathbf{u}(\mathbf{x} + \boldsymbol{\xi}, t) - \mathbf{u}(\mathbf{x}, t)))^2 dV_{\boldsymbol{\xi}} \quad (6)$$

and force vector state

$$\mathbf{T}[\mathbf{x}, t](\boldsymbol{\xi}) = \frac{\omega(|\boldsymbol{\xi}|)}{m} \left[(3K - 5G) \vartheta^{\text{lin}}[\mathbf{x}, t] \boldsymbol{\xi} + 15G \frac{\boldsymbol{\xi} \otimes \boldsymbol{\xi}}{|\boldsymbol{\xi}|^2} (\mathbf{u}(\mathbf{x} + \boldsymbol{\xi}, t) - \mathbf{u}(\mathbf{x}, t)) \right], \quad (7)$$

where the linearized dilatation is given by

$$\vartheta^{\text{lin}}[\mathbf{x}, t] = \frac{3}{m} \int_{\mathcal{H}_x} \omega(|\boldsymbol{\xi}|) \boldsymbol{\xi} \cdot (\mathbf{u}(\mathbf{x} + \boldsymbol{\xi}, t) - \mathbf{u}(\mathbf{x}, t)) dV_{\boldsymbol{\xi}}. \quad (8)$$

The linearized LPS equation of motion is

$$\rho(\mathbf{x}) \frac{\partial^2 \mathbf{u}}{\partial t^2}(\mathbf{x}, t) = \int_{\mathcal{H}_x} \frac{\omega(|\boldsymbol{\xi}|)}{m} \left\{ (3K - 5G) (\vartheta^{\text{lin}}[\mathbf{x}, t] + \vartheta^{\text{lin}}[\mathbf{x} + \boldsymbol{\xi}, t]) \boldsymbol{\xi} + 30G \frac{\boldsymbol{\xi} \otimes \boldsymbol{\xi}}{|\boldsymbol{\xi}|^2} (\mathbf{u}(\mathbf{x} + \boldsymbol{\xi}, t) - \mathbf{u}(\mathbf{x}, t)) \right\} dV_{\boldsymbol{\xi}} + \mathbf{b}(\mathbf{x}, t), \quad (9)$$

and the corresponding equilibrium equation can be expressed as

$$- \int_{\mathcal{H}_x} \frac{\omega(|\boldsymbol{\xi}|)}{m} \left\{ (3K - 5G) (\vartheta^{\text{lin}}[\mathbf{x}] + \vartheta^{\text{lin}}[\mathbf{x} + \boldsymbol{\xi}]) \boldsymbol{\xi} + 30G \frac{\boldsymbol{\xi} \otimes \boldsymbol{\xi}}{|\boldsymbol{\xi}|^2} (\mathbf{u}(\mathbf{x} + \boldsymbol{\xi}) - \mathbf{u}(\mathbf{x})) \right\} dV_{\boldsymbol{\xi}} = \mathbf{b}(\mathbf{x}). \quad (10)$$

Discretized Peridynamic Models

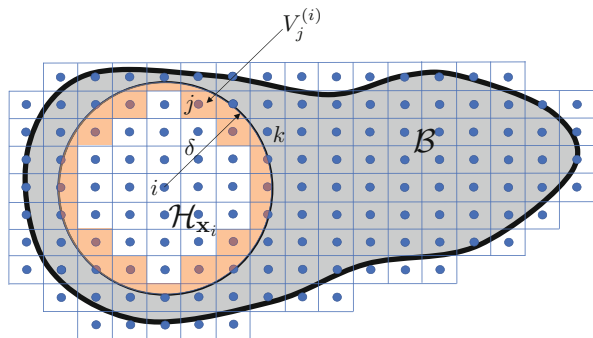
In this chapter, we employ a discretization similar to the meshfree method introduced in Silling and Askari (2005). In that method, no elements or other geometrical connections between computational nodes are used. In our case, we sometimes employ a reference mesh to approximate quadrature weights. However, once these weights are computed, the mesh is discarded and there is no need to computationally track mesh elements in a simulation, in contrast to traditional (mesh-based) finite element methods. For this reason, the methods below are also referred to as meshfree.

To discretize the PD equation of motion, we begin by discretizing the body \mathcal{B} with a set of N^3 nodes $\mathcal{L} = \{\mathbf{x}_i\}_{i=1}^{N^3}$, where each node i is assigned a material volume V_i corresponding to its Voronoi cell, τ_i , such that $\tau_i \cap \tau_j = \emptyset$ for $i \neq j$. The resulting semi-discrete equation is given by

$$\rho_i \frac{d^2 \mathbf{u}_i}{dt^2} = \sum_{j \in \mathcal{F}_i} \{ \mathbf{T}[\mathbf{x}_i, t] \langle \mathbf{x}_j - \mathbf{x}_i \rangle - \mathbf{T}[\mathbf{x}_j, t] \langle \mathbf{x}_i - \mathbf{x}_j \rangle \} V_j^{(i)} + \mathbf{b}_i, \quad (11)$$

where $\rho_i := \rho(\mathbf{x}_i)$, $\mathbf{u}_i := \mathbf{u}(\mathbf{x}_i, t)$, $\mathbf{b}_i := \mathbf{b}(\mathbf{x}_i, t)$, and \mathcal{F}_i is the family of node i , representing the set of all nodes j interacting with node i . Here, the family of a computational node i , \mathcal{F}_i , is a discrete set and should not be confused with the family of a material point \mathbf{x} , $\mathcal{H}_{\mathbf{x}}$, which represents a finite continuum region. In Eq. 11, the quadrature weights $V_j^{(i)}$ represent the volume of the intersection between the cell j and the neighborhood of node i (Seleson 2014). An illustration of the meshfree discretization based on a uniform grid is given in Fig. 1. The quadrature weights $V_j^{(i)}$ may be computed by different numerical algorithms, as described in section “Numerical Tools for Improved Convergence.” When a cell partially overlaps the neighborhood of a given node, e.g., $\mathcal{H}_{\mathbf{x}_i}$ in Fig. 1, the volume of that overlapping region is referred to as a partial volume; these overlaps are illustrated in light orange color in Fig. 1. Algorithms that attempt to approximate

Fig. 1 Meshfree discretization of a peridynamic body \mathcal{B}



such volumes for use as quadrature weights are referred to as *partial-volume algorithms*.

Algorithm 1 : FV (from Silling and Askari (2005))

- 1: {Compute bond length}
- 2: $|\xi| = |\mathbf{x}_j - \mathbf{x}_i|$
- 3: {Check if node j is in the family of node i }
- 4: **if** $|\xi| \leq \delta$ **then**
- 5: $V_j^{(i)} = V_j$
- 6: **else**
- 7: $V_j^{(i)} = 0$
- 8: **end if**
- 9: Return $V_j^{(i)}$

Numerical Tools for Improved Convergence

In the meshfree method proposed in Silling and Askari (2005), the discretization of the PD equation of motion is given by Eq. 11 with a choice of uniform quadrature weights, i.e., $V_j^{(i)} = V_j$ for all $j \in \mathcal{F}_i$, where the family of node i is

$$\mathcal{F}_i = \{j \in \mathcal{L} : |\mathbf{x}_j - \mathbf{x}_i| \leq \delta\}. \quad (12)$$

To easily compare this discretization method with other discretization methods based on partial-volume algorithms, we first present the computation of quadrature weights in this method in Algorithm 1; we refer to it as the FV (Full Volume) algorithm. This discretization method has two drawbacks, as follows:

- (1) **Low spatial integration accuracy:** The interactions between node i and nodes j for which $|\mathbf{x}_j - \mathbf{x}_i| < \delta$ and $\tau_j \not\subset \mathcal{H}_{\mathbf{x}_i}$ are not accurately estimated. Moreover, the interactions between node i and nodes k for which $|\mathbf{x}_k - \mathbf{x}_i| > \delta$ are completely omitted, even when $\tau_k \cap \mathcal{H}_{\mathbf{x}_i} \neq \emptyset$. See Fig. 1 for an illustration. This suggests a low accuracy in the numerical computation of the integral in the PD governing equation, due to a staircase approximation of the boundary of the neighborhood of a given point.
- (2) **Irregular convergence behavior:** A node k for which $|\mathbf{x}_k - \mathbf{x}_i| > \delta$ does not interact with node i . However, under a small grid refinement, such that $\mathbf{x}_k \rightarrow \mathbf{x}_k + \epsilon \hat{\mathbf{e}}_n$ with ϵ a small number and $\hat{\mathbf{e}}_n$ a unit vector so that $|\mathbf{x}_k + \epsilon \hat{\mathbf{e}}_n - \mathbf{x}_i| < \delta$, that node interacts with node i . This suggests that a small grid refinement may cause a significant change in the numerical integration, possibly resulting in abrupt variations in the numerical error. This may result in an irregular (nonsmooth) convergence behavior, as demonstrated in Seleson and Littlewood (2016).

Algorithm 2 : PV-PDLAMMPS (from Parks et al. (2008))

```

1: {Compute bond length}
2:  $|\xi| = |\mathbf{x}_j - \mathbf{x}_i|$ 
3: {Check if cell  $j$  is contained within the neighborhood of node  $i$ 
   (exact in 1D only)}
4: if  $|\xi| + \frac{h}{2} \leq \delta$  then
5:    $V_j^{(i)} = V_j$ 
6: {Check if node  $j$  is in the family of node  $i$ }
7: else if  $|\xi| \leq \delta$  then
8:    $V_j^{(i)} = \frac{1}{h} [\delta - (|\xi| - \frac{h}{2})] V_j$ 
9: else
10:   $V_j^{(i)} = 0$ 
11: end if
12: Return  $V_j^{(i)}$ 

```

A way to mitigate the low integration accuracy and the irregular convergence behavior described above is either to account for all the nodal contributions to the spatial integration in the governing equation, by properly computing all cell-neighborhood intersection volumes, or to weaken the contribution to the spatial integration of nodes near the boundary of the neighborhood of a given node. The former approach has been used by implementing partial-volume algorithms, whereas the latter approach has been used by employing influence functions that smoothly approach zero as the bond length approaches the horizon value (Seleson 2014; Seleson and Littlewood 2016). We describe these approaches in the next sections.

Partial-Volume Algorithms

Partial-volume algorithms are aimed at numerically approximating (or analytically calculating when feasible) volumes of intersections between neighbor cells and the neighborhood of a given node. Those volumes are used as quadrature weights in meshfree discretizations of PD equations. We begin by describing an algorithm proposed in Parks et al. (2008) and implemented in PDLAMMPS (Parks et al. 2011); we refer to it as the PV-PDLAMMPS algorithm (PV is used as an acronym for Partial Volume). This algorithm is based on the family of a given node in Eq. 12 and is described in Algorithm 2. We observe that, as explained in Seleson (2014), the condition to determine whether a cell j is contained within the neighborhood of node i (line 4 of Algorithm 2) is not accurate. Furthermore, the partial-volume correction (line 8 of Algorithm 2) is only a rough approximation to a partial volume based on a simple linear correction.

To improve Algorithm 2, a method was proposed in Hu et al. (2010) and used in Bobaru and Ha (2011). That method is based on the idea of expanding the family of a given node in Eq. 12 to include neighboring nodes located at a distance larger than the horizon from that node with cells intersecting the neighborhood of that node.

The corresponding algorithm is described in Algorithm 3; we refer to it as the PV-HHB algorithm (HHB refers to the initials of the authors of Hu et al. (2010)). It only differs from the PV-PDLAMMPS algorithm in the inequality in line 7.

Although the PV-HHB algorithm represents an improvement to both the FV and PV-PDLAMMPS algorithms, it has the following limitations. First, it does not always account for all neighboring nodes with cells intersecting the neighborhood of a given node, as demonstrated in Seleson (2014). Second, the partial-volume correction is based on the same rough approximation used in the PV-PDLAMMPS algorithm. To correct these issues, a method was proposed in Seleson (2014) for two-dimensional models and expanded to three-dimensional models in Seleson and Littlewood (2016). The idea of that method is to take the family of a given node as all neighboring nodes with cells intersecting the neighborhood of that node, and compute the corresponding intersections as accurately as possible. In Seleson (2014), those intersections were calculated analytically for two-dimensional models discretized over uniform grids, whereas in Seleson and Littlewood (2016) a generalized computational method based on recursive subdivision and sampling was employed to numerically approximate intersection volumes. We refer to the corresponding algorithm as the PV-NC algorithm as in Seleson and Littlewood (2016) (NC is used as an acronym for Numerical Calculation). The algorithm name PV-NC was chosen in Seleson and Littlewood (2016) as a 3D numerically based analog of the 2D algorithm PA-AC (Partial Area-Analytical Calculation), which is based on analytical calculations of cell-neighborhood intersection areas.

Algorithm 3 : PV-HHB (from Hu et al. (2010))

```

1: {Compute bond length}
2:  $|\xi| = |\mathbf{x}_j - \mathbf{x}_i|$ 
3: {Check if cell  $j$  is contained within the neighborhood of node  $i$ 
   (exact in 1D only)}
4: if  $|\xi| + \frac{h}{2} \leq \delta$  then
5:    $V_j^{(i)} = V_j$ 
6: {Check if node  $j$  is in the family of node  $i$ }
7: else if  $|\xi| - \frac{h}{2} \leq \delta$  then
8:    $V_j^{(i)} = \frac{1}{h} [\delta - (|\xi| - \frac{h}{2})] V_j$ 
9: else
10:   $V_j^{(i)} = 0$ 
11: end if
12: Return  $V_j^{(i)}$ 

```

Influence Functions

The computational cost and geometrical challenge posed by the calculation of partial volumes, particularly for nonuniform grids in higher dimensions, motivate the study of a different approach. An alternative method to reduce errors in the

numerical approximation of spatial integrations is to employ kernels that vanish for bonds of a horizon or longer length, particularly smoothly decaying kernels. Such kernels can result from the incorporation of highly regular influence functions. Influence functions in PD models have been originally described in Silling et al. (2007), and their role in peridynamics has been further studied in Seleson and Parks (2011). These functions can be used to select which bonds contribute to the material response of a given point, exclude damaged bonds, and modulate the strength of nonlocal interactions, as well as represent interfaces, free surfaces, and mixtures (Seleson 2010; Silling et al. 2007).

Following Seleson (2014) and Seleson and Littlewood (2016), we employ spherical influence functions of the form

$$\omega(|\xi|) = \begin{cases} \frac{P_n(|\xi|)}{|\xi|^\alpha} & |\xi| \leq \delta, \\ 0 & \text{otherwise,} \end{cases} \quad ; \alpha = 0, 1, \tag{13}$$

where $P_n(r)$ is a polynomial of order $n \in \mathbb{N}_0$, which satisfies $P_n(0) = 1$ and $P_n(\delta) = 0$ for $n > 0$, and $P'_n(r) = P''_n(r) = \dots = P_n^{(k)}(r) = 0$ at $r = 0, \delta$ with $k = (n - 1)/2$ for $n > 1$ odd. Specifically, we have

$$P_0(r) = 1, \tag{14a}$$

$$P_1(r) = 1 - \frac{r}{\delta}, \tag{14b}$$

$$P_3(r) = 1 - 3\left(\frac{r}{\delta}\right)^2 + 2\left(\frac{r}{\delta}\right)^3, \tag{14c}$$

$$P_5(r) = 1 - 10\left(\frac{r}{\delta}\right)^3 + 15\left(\frac{r}{\delta}\right)^4 - 6\left(\frac{r}{\delta}\right)^5, \tag{14d}$$

$$P_7(r) = 1 - 35\left(\frac{r}{\delta}\right)^4 + 84\left(\frac{r}{\delta}\right)^5 - 70\left(\frac{r}{\delta}\right)^6 + 20\left(\frac{r}{\delta}\right)^7. \tag{14e}$$

An illustration of these polynomials is given in Fig. 2.

Convergence Studies of Static Peridynamic Problems

In this section, we present convergence studies of static PD problems following Seleson and Littlewood (2016). Let a body $\mathcal{B} = (0,1) \times (0,1) \times (0,1)$ and let us define a subregion $\Omega = (2\delta, 1-2\delta) \times (2\delta, 1-2\delta) \times (2\delta, 1-2\delta) \subset \mathcal{B}$. All points in Ω are at a distance of at least 2δ from the boundary $\partial\mathcal{B}$, which is equivalent to saying that Ω represents the bulk of the body. The choice of Ω allows us to properly impose displacement boundary conditions, given a boundary function $\mathbf{g}(\mathbf{x})$, in the boundary layer $\overline{\mathcal{B}} \setminus \Omega$ with $\overline{\mathcal{B}} := \mathcal{B} \cup \partial\mathcal{B}$, which contains a 2δ -width volumetric layer required in state-based PD problems. This is in contrast to bond-based PD

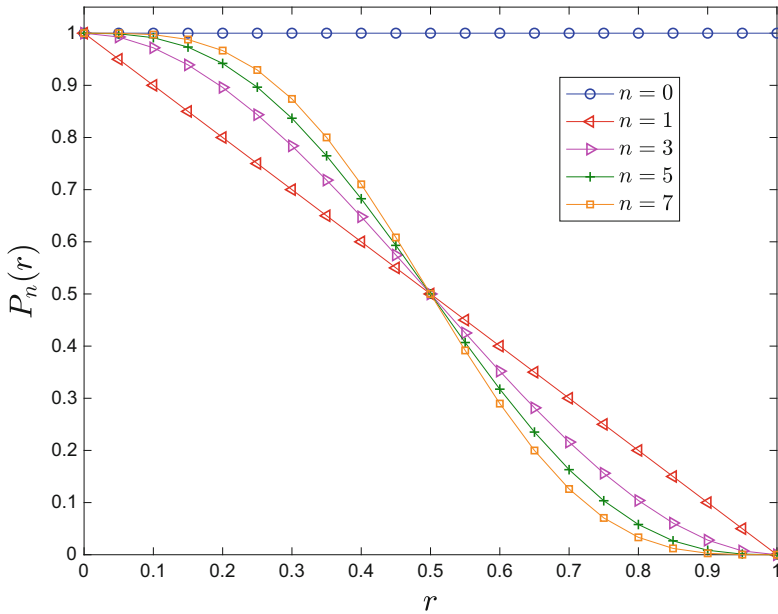


Fig. 2 Illustration of the polynomials $P_n(r)$, $n = 0, 1, 3, 5, 7$, in Eq. 14 for $\delta = 1$

problems, for which a layer of width δ is sufficient. The computational domain and the boundary layer are illustrated in Fig. 3. For simplicity, our static problem employs the linearized LPS model in Eq. 10 and is stated as

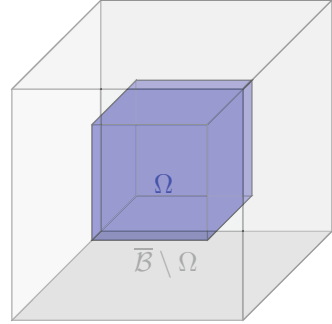
$$\begin{aligned}
 - \int_{\mathcal{H}_x} \frac{\omega(|\xi|)}{m} \left\{ (3K - 5G) (\vartheta^{\text{lin}}[\mathbf{x}] + \vartheta^{\text{lin}}[\mathbf{x} + \xi]) \xi \right. \\
 \left. + 30G \frac{\xi \otimes \xi}{|\xi|^2} (\mathbf{u}(\mathbf{x} + \xi) - \mathbf{u}(\mathbf{x})) \right\} dV_\xi = \mathbf{b}(\mathbf{x}) \quad \mathbf{x} \in \Omega,
 \end{aligned}
 \tag{15a}$$

$$\mathbf{u}(\mathbf{x}) = \mathbf{g}(\mathbf{x}) \quad \mathbf{x} \in \overline{\mathcal{B}} \setminus \Omega.
 \tag{15b}$$

Preliminary Considerations

- (1) *Manufactured solutions:* In this study, we investigate the convergence of numerical solutions of Problem (15) to analytical solutions. For this purpose, we employ the method of manufactured solutions to find a corresponding body force density, given an analytical solution. As proved in Seleson and Littlewood (2016), given a quadratic displacement field, the PD equilibrium Eq. 10 reduces to the classical Navier-Cauchy equation of static elasticity,

Fig. 3 Illustration of the computation domain Ω and boundary layer $\bar{B} \setminus \Omega$ for Problem (15)



$$-\left[G \nabla^2 \mathbf{u}(\mathbf{x}) + \left(K + \frac{1}{3} G \right) \nabla (\nabla \cdot \mathbf{u})(\mathbf{x}) \right] = \mathbf{b}(\mathbf{x}). \tag{16}$$

Consequently, we can employ Eq. 16 to find a corresponding body force density for our PD problem as follows. Let a displacement field $\mathbf{u} = (u, v, w)$, where

$$u(\mathbf{x}) = U_{11}x^2 + U_{22}y^2 + U_{33}z^2 + U_{12}xy + U_{13}xz + U_{23}yz, \tag{17}$$

$$v(\mathbf{x}) = V_{11}x^2 + V_{22}y^2 + V_{33}z^2 + V_{12}xy + V_{13}xz + V_{23}yz, \tag{18}$$

$$w(\mathbf{x}) = W_{11}x^2 + W_{22}y^2 + W_{33}z^2 + W_{12}xy + W_{13}xz + W_{23}yz, \tag{19}$$

where $\mathbf{x} = (x, y, z)$, and $U_{11}, U_{22}, U_{33}, U_{12}, U_{13}, U_{23}, V_{11}, V_{22}, V_{33}, V_{12}, V_{13}, V_{23}, W_{11}, W_{22}, W_{33}, W_{12}, W_{13},$ and W_{23} are constant coefficients. Then, the components of the body force density $\mathbf{b} = (b_1, b_2, b_3)$ are:

$$b_1 = -\left[2G (U_{11} + U_{22} + U_{33}) + \left(K + \frac{1}{3} G \right) (2U_{11} + V_{12} + W_{13}) \right],$$

$$b_2 = -\left[2G (V_{11} + V_{22} + V_{33}) + \left(K + \frac{1}{3} G \right) (U_{12} + 2V_{22} + W_{23}) \right],$$

$$b_3 = -\left[2G (W_{11} + W_{22} + W_{33}) + \left(K + \frac{1}{3} G \right) (U_{13} + V_{23} + 2W_{33}) \right].$$

(2) *Error norm for numerical solutions:* Let the body \mathcal{B} be discretized with a set of N^3 nodes $\mathcal{L} = \{\mathbf{x}_i\}_{i=1}^{N^3}$ over a uniform grid with grid spacing h , so that the Voronoi cell of node i is the cubic cell $\tau_i = \left(x_i - \frac{h}{2}, x_i + \frac{h}{2} \right) \times \left(y_i - \frac{h}{2}, y_i + \frac{h}{2} \right) \times \left(z_i - \frac{h}{2}, z_i + \frac{h}{2} \right)$, where $\mathbf{x}_i = (x_i, y_i, z_i)$. In particular, in the numerical simulations we employ a discretization with N nodes per direction, where each computational node possesses a full cubic cell inside the (cubic) body, i.e., all nodes are at a distance of a least $h/2$ from the boundary of

the body. To compute an error norm, we assume the numerical solution to be given by a piecewise-constant approximation $\mathbf{u}^h = (u^h, v^h, w^h)$, so that $\mathbf{u}^h(\mathbf{x}) = (u_i^h, v_i^h, w_i^h)$ for all $\mathbf{x} \in \tau_i$ with u_i^h, v_i^h , and w_i^h constants. The L^2 -norm of the error is then

$$\begin{aligned} \|\mathbf{u}^h - \mathbf{u}\|_2 &= \left[\int_{\mathcal{B}} (u^h(\mathbf{x}) - u(\mathbf{x}))^2 + (v^h(\mathbf{x}) - v(\mathbf{x}))^2 + (w^h(\mathbf{x}) - w(\mathbf{x}))^2 dV_{\mathbf{x}} \right]^{\frac{1}{2}} \\ &= \left[\sum_{i=1}^{N^3} \int_{\tau_i} (u_i^h - u(\mathbf{x}))^2 + (v_i^h - v(\mathbf{x}))^2 + (w_i^h - w(\mathbf{x}))^2 dV_{\mathbf{x}} \right]^{\frac{1}{2}} \\ &= \left[\sum_{i=1}^{N^3} \sum_{g=1}^{N_G} \sum_{\ell=1}^{N_L} \sum_{s=1}^{N_S} \left\{ \left(u_i^h - u \left(x_i + \frac{h}{2} \alpha_g, y_i + \frac{h}{2} \beta_{\ell}, z_i + \frac{h}{2} \gamma_s \right) \right)^2 \right. \right. \\ &\quad \left. \left. + \left(v_i^h - v \left(x_i + \frac{h}{2} \alpha_g, y_i + \frac{h}{2} \beta_{\ell}, z_i + \frac{h}{2} \gamma_s \right) \right)^2 \right. \right. \\ &\quad \left. \left. + \left(w_i^h - w \left(x_i + \frac{h}{2} \alpha_g, y_i + \frac{h}{2} \beta_{\ell}, z_i + \frac{h}{2} \gamma_s \right) \right)^2 \right\} \frac{h^3}{8} w_g \widehat{w}_{\ell} \widetilde{w}_s \right]^{\frac{1}{2}}, \end{aligned} \tag{20}$$

where $\{\alpha_g\}_{g=1, \dots, N_G}$ and $\{w_g\}_{g=1, \dots, N_G}$; $\{\beta_{\ell}\}_{\ell=1, \dots, N_L}$ and $\{\widehat{w}_{\ell}\}_{\ell=1, \dots, N_L}$; and $\{\gamma_s\}_{s=1, \dots, N_S}$ and $\{\widetilde{w}_s\}_{s=1, \dots, N_S}$ are Gauss quadrature points and weights in the standard element $[-1, 1]$ along the x -, y -, and z -directions, respectively. A quadrature rule with N_G, N_L , and N_S Gauss quadrature points along the x -, y -, and z -directions, respectively, integrates exactly monomials of the form $x^g y^{\ell} z^s$ with $0 \leq g \leq 2N_G - 1, 0 \leq \ell \leq 2N_L - 1$, and $0 \leq s \leq 2N_S - 1$. Since the analytical solution is assumed quadratic and we integrate a function of the solution squared, we would like to integrate exactly monomials up to $g = 4, \ell = 4$, and $s = 4$. We thus choose $N_G = N_L = N_S = 3$.

- (3) *Computational cost and horizon choice:* In this type of convergence study, the horizon needs to be carefully chosen such that the size of the boundary layer is small relative to the size of the computational domain, yet the computational cost remains tractable. To clarify this, let a body \mathcal{B} be a cube of edge length L and the computational domain Ω be a smaller cube of edge length $L - 4\delta$ (as illustrated in Fig. 3), and define $N_{\text{neig}} = \delta/h$ with h the grid spacing. Then, the total number of computational nodes required to discretize the body \mathcal{B} is

$$N^3 = \left(\frac{L}{h} \right)^3 = \left(N_{\text{neig}} \frac{L}{\delta} \right)^3. \tag{21}$$

Table 1 Total number of computational nodes

$N_{\text{neig}} (\delta/h)$	$\delta = 0.10L$	$\delta = 0.04L$	$\delta = 0.02L$	$\delta = 0.01L$
1	1,000	15,625	125,000	1,000,000
2	8,000	125,000	1,000,000	8,000,000
3	27,000	421,875	3,375,000	27,000,000
4	64,000	1,000,000	8,000,000	64,000,000
5	125,000	1,953,125	15,625,000	125,000,000
6	216,000	3,375,000	27,000,000	216,000,000

In Table 1, we show the total number of computational nodes required to discretize \mathcal{B} , for various choices of horizon δ . We observe that the horizon cannot be too small, because otherwise the total number of computational nodes may exceed the available computational resources. For instance, for $\delta = 0.01L$, a computation with $N_{\text{neig}} = 6$ would require 216 million nodes.

We now demonstrate that the horizon cannot be too large either. For this purpose, we look at the ratio of the boundary layer to the computational domain:

$$\frac{|\mathcal{B} \setminus \Omega|}{|\Omega|} = \frac{L^3 - (L - 4\delta)^3}{(L - 4\delta)^3} = \frac{1}{(1 - 4\delta/L)^3} - 1, \quad (22)$$

where $|\Omega|$ denotes the volume of the domain Ω . This ratio rapidly grows with increasing ratio δ/L , as shown in Fig. 4. In particular, for $\delta > 0.0516L$, the volume of the boundary layer $\mathcal{B} \setminus \Omega$ exceeds that of the computational domain Ω . To exemplify this, specific values of Eq. 22 are shown in Table 2.

- (4) *Convergence avenue and refinement path:* As described in the Introduction, there exist different types of convergence in peridynamics. Here, we would like to study the convergence of numerical solutions of static PD problems to the corresponding analytical PD solutions. For this purpose, we fix the horizon δ and employ a grid refinement. This, in fact, is analogous to standard convergence studies in numerical partial differential equations (PDEs), where the grid (or mesh) is refined. Since the horizon is fixed, the ratio δ/h increases under grid refinement.

As demonstrated in Table 1, we cannot afford to refine the mesh beyond a certain ratio $N_{\text{neig}} = \delta/h$ because the number of required computational nodes would become intractable. On the other hand, to produce meaningful comparisons in our study, we need enough convergence data points. For the purpose of our studies, we took $L = 1$ and $\delta = 0.04$, and we performed a refinement given by $N_{\text{neig}} = 3, 4, 5, 6$. The data used here is a subset of the results presented in Seleson and Littlewood (2016). The reader should refer to that reference for a refinement study with additional data points.

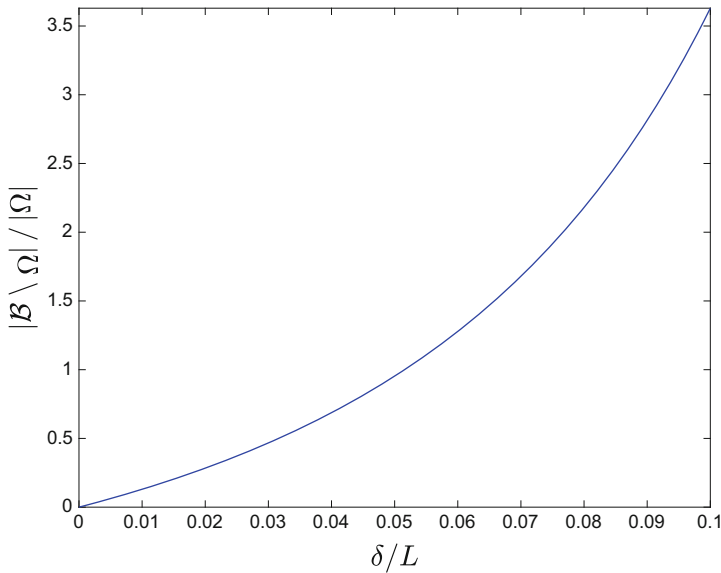


Fig. 4 Ratio $|\mathcal{B} \setminus \Omega|/|\Omega|$ as a function of the horizon δ in a cubic body \mathcal{B} of edge length L with a cubic subdomain Ω of edge length $L - 4\delta$

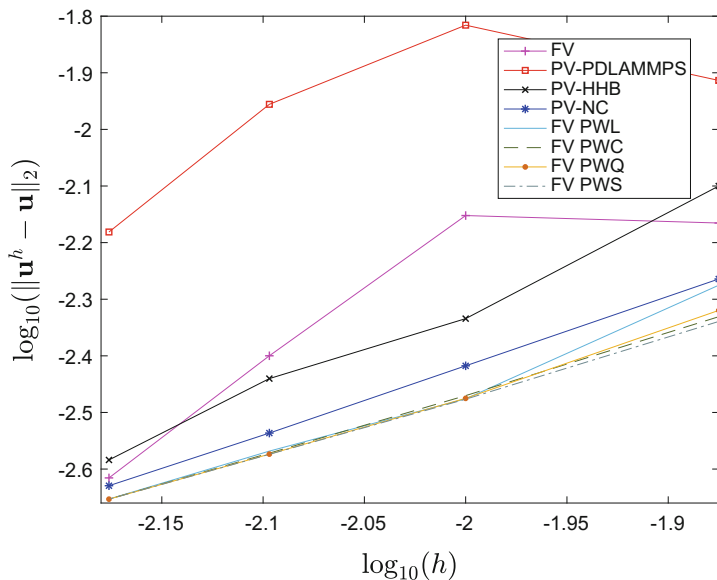
Table 2 Volume ratio $|\mathcal{B} \setminus \Omega|/|\Omega|$

δ/L	0.000	0.001	0.005	0.010	0.050	0.100
$ \mathcal{B} \setminus \Omega / \Omega $	0.000	0.012	0.062	0.130	0.953	3.630

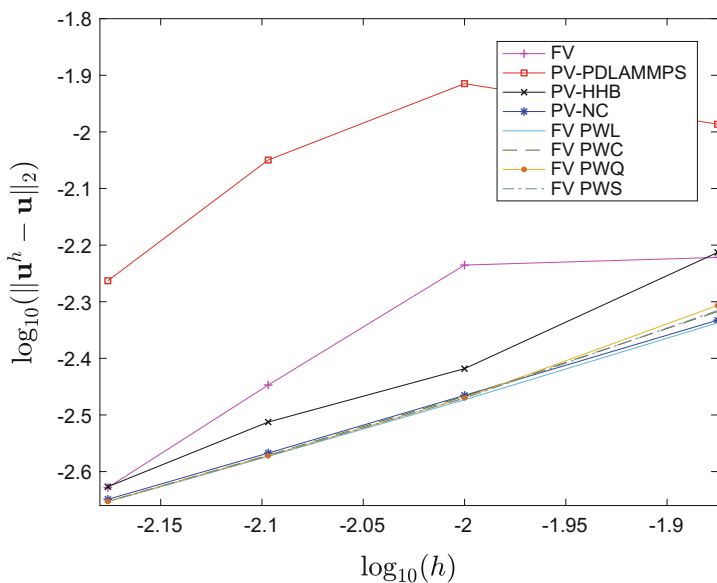
Numerical Results

We study the convergence of numerical solutions of Problem (15) with the following choices: $\delta = 0.04$, $K = 1$, $G = 0.5$, $\mathbf{b} = (-\frac{10}{3}, 0, 0)$, and $\mathbf{g}(\mathbf{x}) = (x^2, 0, 0)$, and where the weighted volume m is computed analytically. The chosen elastic moduli correspond to a material with Poisson’s ratio $\nu = 0.29$. The goal of this study is, on one hand, to compare the performance of the FV algorithm with the partial-volume algorithms PV-PDLAMMPS, PV-HHB, and PV-NC, and, on the other hand, to study the effect of incorporating specific influence functions within the PD force vector state field. The total number of computational nodes employed in the simulations appears in the third column of Table 1. The simulations were run in *Peridigm*, a PD code developed at Sandia National Laboratories (Parks et al. 2012); see Seleson and Littlewood (2016) for details regarding the solution procedure.

In Fig. 5, we present the convergence results for the FV algorithm as well as for the different partial-volume algorithms and influence functions for (a) $\alpha = 0$ and (b) $\alpha = 1$, where the error is computed with the L^2 -norm given in Eq. 20. Specifically, we examine influence functions of the following form: piecewise linear (PWL) with $P_1(r)$ in Eq. 14b; piecewise cubic (PWC) with $P_3(r)$ in Eq. 14c;



(a) $\alpha = 0$



(b) $\alpha = 1$

Fig. 5 Convergence of the numerical solution of Problem (15) using different partial-volume algorithms and different influence functions

Table 3 Convergence rates for the (static) convergence results in Fig. 5

Algorithm	$\alpha = 0$		$\alpha = 1$	
	\bar{r}	R	\bar{r}	R
FV	1.53	0.165	1.38	0.128
PV-PDLAMMPS	0.86	0.186	0.89	0.167
PV-HHB	1.56	0.035	1.34	0.030
PV-NC	1.22	0.003	1.05	0.001
FV PWL	1.24	0.036	1.05	0.004
FV PWC	1.07	0.005	1.11	0.009
FV PWQ	1.10	0.014	1.15	0.016
FV PWS	1.04	0.006	1.12	0.012

piecewise quintic (PWQ) with $P_5(r)$ in Eq. 14d; and piecewise septic (PWS) with $P_7(r)$ in Eq. 14e. In order to estimate a convergence rate, to each curve we fit a linear function (in a least-squares sense) of the form

$$f(h) = \bar{r} \log_{10}(h) + q, \quad (23)$$

where \bar{r} and q are constant coefficients. The value of \bar{r} represents an “average” convergence rate. To estimate the quality of the linear fit, we also compute the norm of residuals,

$$R = \sqrt{\sum_{n=1}^{N_{\text{iter}}} (f(h_n) - \log_{10}(E_n))^2}, \quad (24)$$

where E_n and h_n are the L^2 -norm of the error (cf. Eq. 20) and the grid spacing, respectively, at the refinement step n , and N_{iter} is the total number of refinement steps. These results are presented in Table 3. We observe that both methods, accurate estimation of partial volumes (see PV-NC) and employment of smooth influence functions (see, e.g., FV PWS), provide a significant improvement relative to the FV algorithm, as clearly observed in the reduction in the value of R . Based on the convergence rates obtained, we conclude that those methods approach a first-order convergence.

Convergence Studies of Dynamic Peridynamic Problems

In this section, we present convergence studies of dynamic PD problems, based on unconstrained wave propagation. As in section “Convergence Studies of Static Peridynamic Problems,” we investigate the effect of different partial-volume algorithms and influence functions on the convergence behavior, as the grid is refined under a fixed horizon.

Let a body $\mathcal{B} = (-\frac{1}{2}, \frac{1}{2}) \times (-\frac{1}{2}, \frac{1}{2}) \times (-\frac{1}{2}, \frac{1}{2})$. Our dynamic problem employs the linearized LPS model in Eq. 9 and is stated as

$$\rho(\mathbf{x}) \frac{\partial^2 \mathbf{u}}{\partial t^2}(\mathbf{x}, t) = \int_{\mathcal{H}_{\mathbf{x}}} \frac{\omega(|\boldsymbol{\xi}|)}{m} \left\{ (3K - 5G) (\vartheta^{\text{lin}}[\mathbf{x}, t] + \vartheta^{\text{lin}}[\mathbf{x} + \boldsymbol{\xi}, t]) \boldsymbol{\xi} + 30G \frac{\boldsymbol{\xi} \otimes \boldsymbol{\xi}}{|\boldsymbol{\xi}|^2} (\mathbf{u}(\mathbf{x} + \boldsymbol{\xi}, t) - \mathbf{u}(\mathbf{x}, t)) \right\} dV_{\boldsymbol{\xi}} + \mathbf{b}(\mathbf{x}, t) \quad \mathbf{x} \in \bar{B}, t \geq 0, \tag{25a}$$

$$\mathbf{u}(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x}) \quad \mathbf{x} \in \bar{B}, \tag{25b}$$

$$\frac{\partial \mathbf{u}}{\partial t}(\mathbf{x}, 0) = \mathbf{v}_0(\mathbf{x}) \quad \mathbf{x} \in \bar{B}, \tag{25c}$$

where \mathbf{u}_0 and \mathbf{v}_0 are displacement and velocity initial conditions. As opposed to the static Problem (15), no displacement boundary conditions are imposed in these dynamic studies, eliminating the need to define a boundary layer.

Preliminary Considerations

(1) *Error norm for numerical solutions:* The error for the dynamic PD problems was computed by comparing a series of increasingly refined discretizations against results obtained using a highly refined reference discretization. This was necessitated by the lack of analytical solutions of the state-based PD Problem (25). While analytical solutions for wave propagation concerning linear elastic bond-based PD models have been presented in one dimension in Weckner and Abeyaratne (2005) and in higher dimensions in Weckner et al. (2009), analytical solutions are not currently available for the state-based PD Problem (25). Unfortunately, our approach for computing the error is extremely computationally demanding, because the number of PD bonds increases dramatically as the mesh is refined under a fixed horizon, as demonstrated in Table 4. To achieve numerical accuracy while remaining computationally tractable, a value of $N_{\text{neig}} = 10$ was employed for the reference numerical solution.

The calculation of an error norm in the dynamic studies followed assumptions analogous to those in the static simulations. We assumed the numerical

Table 4 Total number of peridynamic (PD) bonds in a simulation for $\delta = 0.05$

$N_{\text{neig}} (\delta/h)$	Number of PD bonds	
	FV	PV-NC
3	12,433,244	25,077,672
4	62,022,592	110,046,364
5	242,986,412	384,681,876
6	753,964,092	1,040,684,328
7	1,838,660,296	2,552,461,732
8	4,080,378,204	5,479,353,788
9	8,456,684,628	10,782,968,496
10	15,752,838,172	19,683,573,672

solution for both the trial simulations and the reference simulation to be given by a piecewise-constant approximation $\mathbf{u}^h = (u^h, v^h, w^h)$, so that $\mathbf{u}^h(\mathbf{x}, t) = (u_i^h, v_i^h, w_i^h)$ for all $\mathbf{x} \in \tau_i$ at time t with u_i^h, v_i^h , and w_i^h constants. We denote by τ_i^{ref} a Voronoi cell in the reference discretization. To avoid having the reference discretization being an exact subdivision of any of the trial discretizations, which could result in biased error calculations for those specific discretizations, the grid in the reference discretization was perturbed so that the internal nodes (i.e., nodes not directly adjacent to the boundary of the body) were shifted a distance of 0.3 times the reference grid spacing, h_{ref} , in each direction. This perturbation does not affect the error calculation in our studies because in the wave propagation problems considered here nodes adjacent to the boundary of the body have a zero displacement over the course of the simulations. Let the domain \mathcal{B} be decomposed into N^{int} subdomains $\{\tilde{\tau}_n\}_{n=1, \dots, N^{\text{int}}}$, so that $\tilde{\tau}_n = \tau_i \cap \tau_j^{\text{ref}}$ for given i and j ; in other words, these subdomains are intersections of pairs of Voronoi cells, one from the trial discretization and one from the reference discretization. Then, the L^2 -norm of the error for a given level of grid refinement, given by a grid spacing h , at the final time T can be computed as

$$\begin{aligned} \|\mathbf{u}^h - \mathbf{u}^{h_{\text{ref}}}\|_2 &= \left[\int_{\mathcal{B}} (u^h(\mathbf{x}, T) - u^{h_{\text{ref}}}(\mathbf{x}, T))^2 + (v^h(\mathbf{x}, T) - v^{h_{\text{ref}}}(\mathbf{x}, T))^2 \right. \\ &\quad \left. + (w^h(\mathbf{x}, T) - w^{h_{\text{ref}}}(\mathbf{x}, T))^2 dV_{\mathbf{x}} \right]^{\frac{1}{2}} \\ &= \left[\sum_{n=1}^{N^{\text{int}}} \left((u_n^h - u_n^{h_{\text{ref}}})^2 + (v_n^h - v_n^{h_{\text{ref}}})^2 + (w_n^h - w_n^{h_{\text{ref}}})^2 \right) V_n^{\text{int}} \right]^{\frac{1}{2}}, \end{aligned} \quad (26)$$

where (u_n^h, v_n^h, w_n^h) and $(u_n^{h_{\text{ref}}}, v_n^{h_{\text{ref}}}, w_n^{h_{\text{ref}}})$ are the trial and reference numerical solutions, respectively, over the subdomain $\tilde{\tau}_n$ at the final time T , and V_n^{int} is the volume of that subdomain.

- (2) *Reference solution:* The solution of the PD Problem (25) varies for different choices of influence function. This is true in general for PD problems, as demonstrated in Seleson and Parks (2011), but was not the case for the static simulations in section “[Convergence Studies of Static Peridynamic Problems](#)” due to the particular choice of PD model and applied body force density. On the other hand, even when the influence function is unchanged, the use of different partial-volume algorithms may produce different numerical solutions. As the reference solution is based on a (highly refined) numerical solution, and not on an analytical one, separate reference solutions were computed for the dynamic PD problems, one for each of the partial-volume algorithms and influence functions.

- (3) *Simulation running time*: In these dynamic simulations, we would like to investigate the accuracy of the numerical results for various choices of partial-volume algorithms and influence functions. This suggests the need to let the solution evolve sufficiently such that the effect of the discretization on the system dynamics is clearly observable. On the other hand, we do not want to concern ourselves with wave reflections from domain boundaries or with interactions of a propagating wave with itself, as this may introduce another level of complexity in the dynamics of the system. As a result, we run simulations based on an initial pulse concentrated away from the center of the domain and away from the domain boundaries, and we allow the pulse to propagate as much as possible without interacting with the domain boundaries or with itself at the center of the domain.
- (4) *Signal resolution and discretization*: As discussed in section “[Preliminary Considerations](#)” of the static PD problems, extremely refined discretizations must be avoided due to computational limitations. However, in the dynamic simulations, the discretization cannot be too coarse either because we need to resolve the shape of the propagating wave in the domain.

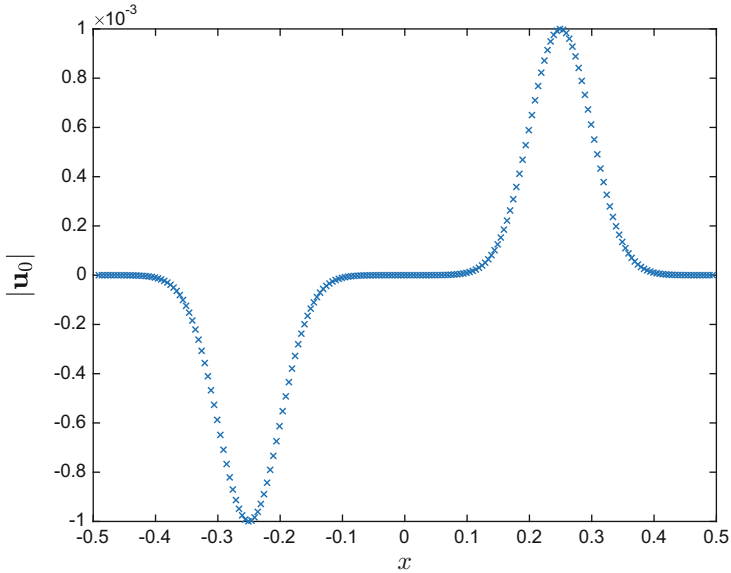
Numerical Results

We study the convergence of numerical solutions of Problem (25) with the following choices: $\delta = 0.05$, $K = 1$, $G = 0.5$, $\mathbf{b} = \mathbf{0}$, and $\rho = 100.0$; as in the static Problem (15), the weighted volume m is computed analytically. The initial conditions are given as follows:

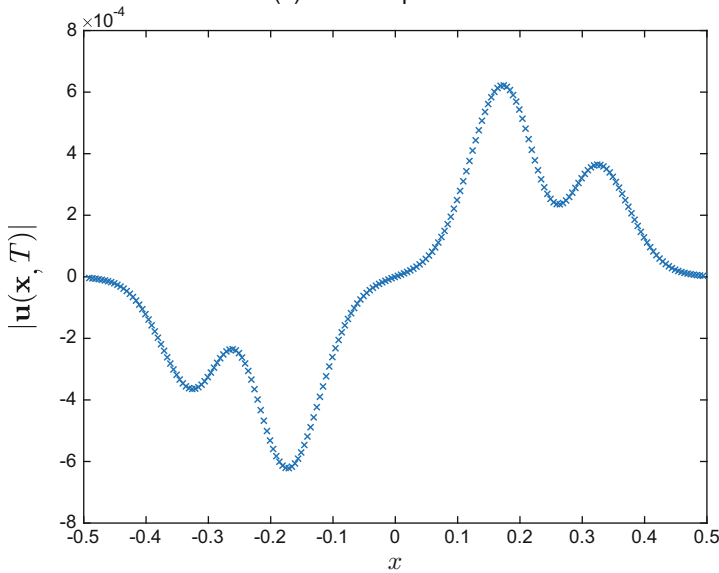
$$\mathbf{u}_0(\mathbf{x}) = \begin{cases} ae^{-\frac{(|\mathbf{x}|-r_0)^2}{\ell^2}} \frac{\mathbf{x}}{|\mathbf{x}|} & \text{if } (r_0 - 3\ell) \leq |\mathbf{x}| \leq (r_0 + 3\ell), \\ \mathbf{0} & \text{otherwise,} \end{cases} \quad (27)$$

$$\mathbf{v}_0(\mathbf{x}) = \mathbf{0}, \quad (28)$$

where the constants were assigned values of $a = 0.001$, $r_0 = 0.25$, and $\ell = 0.07$. The initial displacement field is a spherical layer, with a Gaussian radial distribution. The parameter r_0 represents the center of the Gaussian and the parameter ℓ determines its width. Points in that layer are assigned an initial outward radial displacement as determined by the Gaussian function, whereas points outside that layer are given a zero initial displacement. The width parameter ℓ was chosen so that the coarsest discretization can resolve the shape of the Gaussian distribution. In addition, we observe that at a distance $r = 3\ell$ from the center of the Gaussian, the amplitude of the initial displacement decays to about $1.2E-7$, suggesting a reasonable transition to the zero displacement field. No initial velocity is given. Plots of the initial displacement field are given in Figs. 6a and 7a. The wave is allowed to propagate freely from the initial time $t_0 = 0$ to the final time $T = 0.6$. Time integration was achieved using the standard velocity-Verlet scheme with a fixed time step $\Delta t = 0.01$,



(a) Initial displacement



(b) Final displacement

Fig. 6 Initial displacement (a) and final displacement (b) for nodes located along the x -axis for the three-dimensional wave propagation, based on the reference finest grid with grid spacing $h_{ref} = 0.005$ ($\delta/h_{ref} = 10$). The final displacement was computed with the FV algorithm for $\alpha = 0$

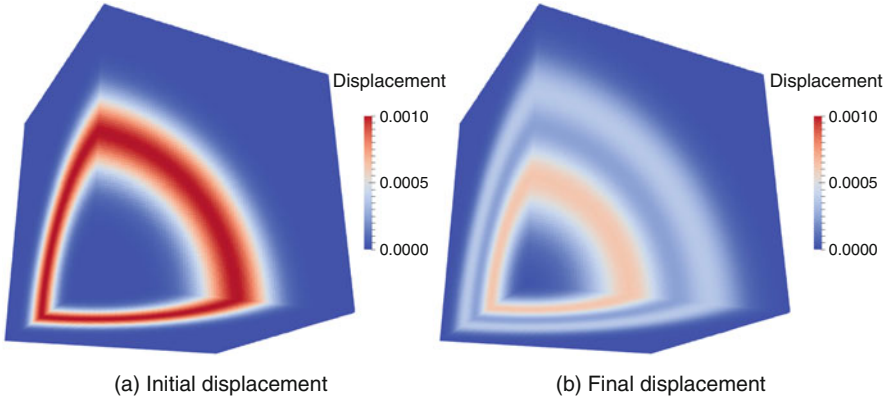
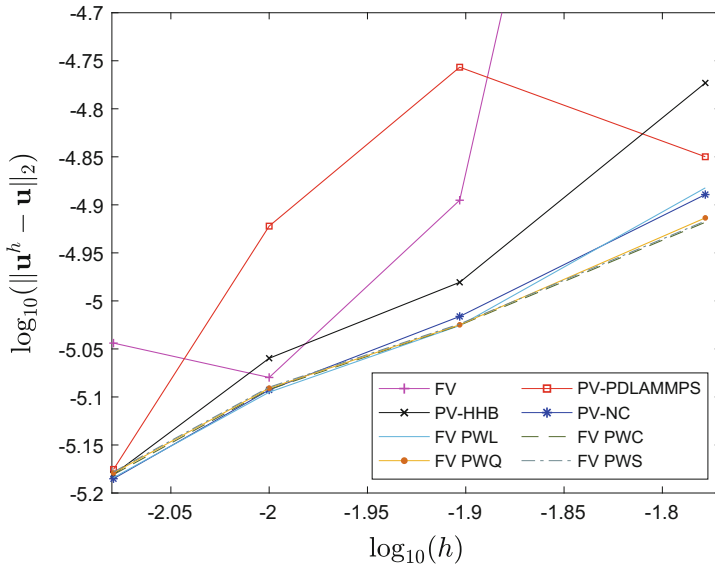


Fig. 7 Initial displacement (a) and final displacement (b) on an octant of the cubic domain for the three-dimensional wave propagation, based on the reference finest grid with grid spacing $h_{\text{ref}} = 0.005$ ($\delta/h_{\text{ref}} = 10$). The final displacement was computed with the FV algorithm for $\alpha = 0$. Coloration denotes displacement magnitude

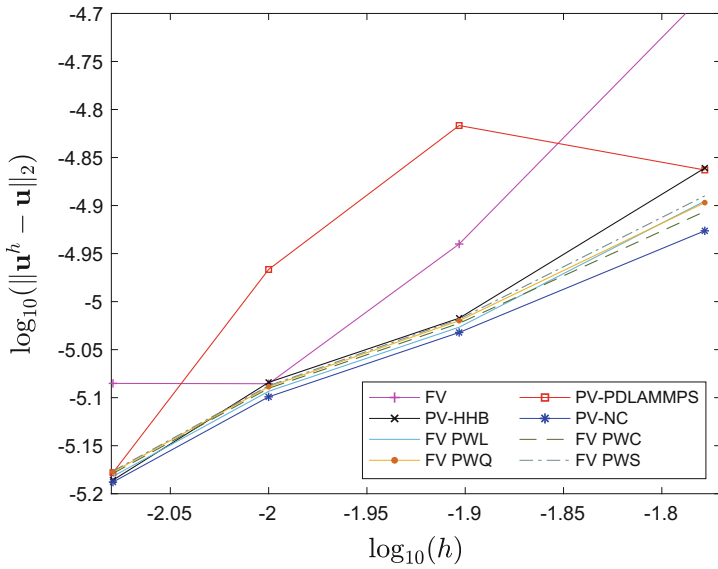
which was stable for all the simulations. The final displacement field numerically computed with the FV algorithm for $\alpha = 0$ is given in Figs. 6b and 7b.

In Fig. 8, we present the convergence results for the different partial-volume algorithms and influence functions for (a) $\alpha = 0$ and (b) $\alpha = 1$, where the error is computed with the L^2 -norm given by Eq. 26. As in the static Problem (15), to each curve we fit a linear function of the form of Eq. 23 to estimate a convergence rate; to also estimate the quality of the linear fit, we compute the norm of residuals with Eq. 24. These results are presented in Table 5. As in the static problems, both methods, accurate estimation of partial volumes (see PV-NC) and employment of smooth influence functions (see, e.g., FV PWS), provide a significant improvement relative to the FV algorithm, as clearly observed in the reduction in the value of R . Based on the convergence rates obtained, we conclude that those methods approach a first-order convergence.

To demonstrate the effect of the influence function on the solution of the dynamic problem, we compare in Fig. 9 the final displacement field for points along the x -axis, for the different partial-volume algorithms and influence functions. The full displacement field is presented in the top plots for (a) $\alpha = 0$ and (b) $\alpha = 1$. To emphasize the differences between the two numerical tools discussed in this chapter, partial-volume algorithms and influence functions, in the bottom plots, we zoom in to the positive x -axis and simply compare the FV algorithm to the PV-NC and FV PWS algorithms for (c) $\alpha = 0$ and (d) $\alpha = 1$. The results suggest that the solution for the smooth influence function is different than the one obtained with the FV algorithm; on the other hand, using a partial-volume correction seems to more closely preserve the dynamics of the original model. This is natural, since partial-volume algorithms are only aimed at improving the discretization of the PD governing equation, without modifying the continuum model. In contrast, the



(a) $\alpha = 0$



(b) $\alpha = 1$

Fig. 8 Convergence of the numerical solution of Problem (25) at the final time T using different partial-volume algorithms and different influence functions

Table 5 Convergence rates for the (dynamic) convergence results in Fig. 8

Algorithm	$\alpha = 0$		$\alpha = 1$	
	\bar{r}	R	\bar{r}	R
FV	4.27	0.514	1.41	0.099
PV-PDLAMMPS	1.05	0.202	1.02	0.157
PV-HHB	1.31	0.038	1.04	0.026
PV-NC	0.96	0.013	0.85	0.016
FV PWL	0.98	0.019	0.93	0.017
FV PWC	0.85	0.016	0.88	0.015
FV PWQ	0.86	0.015	0.91	0.015
FV PWS	0.85	0.016	0.93	0.015

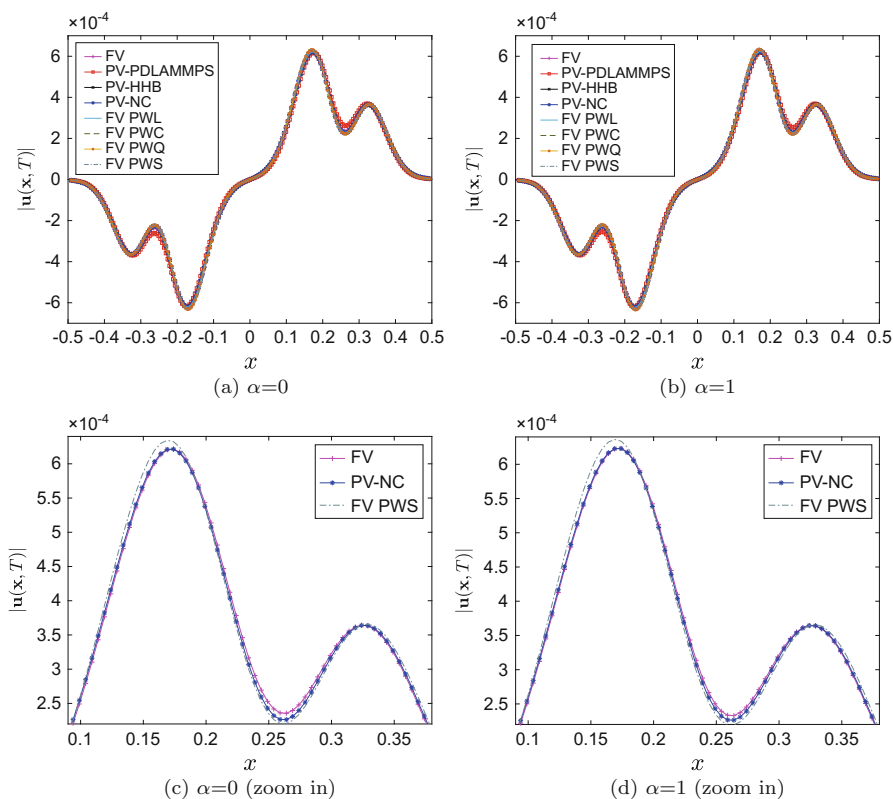


Fig. 9 Final displacement for nodes located along the x -axis for the three-dimensional wave propagation, for different partial-volume algorithms and influence functions, based on the reference finest grid with grid spacing $h_{ref} = 0.005$ ($\delta/h_{ref} = 10$). In the top plots, the entire displacement field is compared for (a) $\alpha = 0$ and (b) $\alpha = 1$. In the bottom plots, a zoom in to a portion of the curve along the positive x -axis is presented comparing the FV, PV-NC, and FV PWS algorithms for (c) $\alpha = 0$ and (d) $\alpha = 1$

incorporation of different influence functions alters the continuum model itself. This is consistent with the results presented in Seleson and Parks (2011). Further dynamic studies comparing the use of partial-volume algorithms and smooth influence functions are currently underway and will be reported in the future.

Concluding Remarks

This chapter considered convergence studies for linear elastic state-based peridynamic models, based on meshfree discretizations, and presented two numerical tools, partial-volume algorithms and influence functions, to improve the convergence behavior of numerical solutions of static and dynamic peridynamic problems. The work in this chapter drew from, and extended, the convergence studies presented for static peridynamic problems in Seleson and Littlewood (2016). We observed that both numerical tools can assist in improving the convergence behavior of static and dynamic peridynamic simulations, resulting in approximately first-order convergence. Depending on the application and problem under consideration, influence functions can, however, significantly affect the dynamical behavior of peridynamic systems.

An important contribution of this chapter is a discussion regarding various challenges present in convergence studies in peridynamic problems, many of which are not found in analogous convergence studies for PDEs. We hope that such discussion will serve as guidance for future related research studies.

Acknowledgments This study was supported in part by the Householder Fellowship which is jointly funded by: the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Applied Mathematics program, under award number ERKJE45, and the Laboratory Directed Research and Development program at the Oak Ridge National Laboratory (ORNL), which is operated by UT-Battelle, LLC., for the U.S. Department of Energy under Contract DE-AC05-00OR22725; and by the Laboratory Directed Research and Development program at Sandia National Laboratories, which is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

References

- E. Askari, F. Bobaru, R.B. Lehoucq, M.L. Parks, S.A. Silling, O. Weckner, Peridynamics for multiscale materials modeling, in *SciDAC 2008, 13–17 July, Seattle*, Journal of Physics: Conference Series, vol. 125, IOP Publishing, 2008. 012078
- F. Bobaru, Y.D. Ha, Adaptive refinement and multiscale modeling in 2D peridynamics. *Int. J. Multiscale Comput. Eng.* **9**, 635–659 (2011)
- F. Bobaru, Y.D. Ha, W. Hu, Damage progression from impact in layered glass modeled with peridynamics. *Centr. Eur. J. Eng.* **2**, 551–561 (2012)
- F. Bobaru, M. Yang, L.F. Alves, S.A. Silling, E. Askari, J. Xu, Convergence, adaptive refinement, and scaling in 1D peridynamics. *Int. J. Numer. Meth. Eng.* **77**, 852–877 (2009)

- X. Chen, M. Gunzburger, Continuous and discontinuous finite element methods for a peridynamics model of mechanics. *Comput. Methods Appl. Mech. Eng.* **200**, 1237–1250 (2011)
- D. De Meo, N. Zhu, E. Oterkus, Peridynamic modeling of granular fracture in polycrystalline materials. *J. Eng. Mater. Technol.* **138**, 041008–041008–16 (2016)
- Q. Du, X. Tian, Asymptotically compatible schemes for peridynamics based on numerical quadratures, in *Proceedings of the ASME 2014 International Mechanical Engineering Congress and Exposition*, Montreal, 2014. Paper No. IMECE2014-39620
- W. Gerstle, N. Sau, S. Silling, Peridynamic modeling of concrete structures. *Nucl. Eng. Des.* **237**, 1250–1258 (2007)
- M. Ghajari, L. Iannucci, P. Curtis, A peridynamic material model for the analysis of dynamic crack propagation in orthotropic media. *Comput. Methods Appl. Mech. Eng.* **276**, 431–452 (2014)
- Y.D. Ha, F. Bobaru, Studies of dynamic crack propagation and crack branching with peridynamics. *Int. J. Fract.* **162**, 229–244 (2010)
- Y.D. Ha, F. Bobaru, Characteristics of dynamic brittle fracture captured with peridynamics. *Eng. Fract. Mech.* **78**, 1156–1168 (2011)
- W. Hu, Y. D. Ha, F. Bobaru, Numerical integration in peridynamics, Technical report, University of Nebraska-Lincoln, Sept. 2010
- W. Hu, Y.D. Ha, F. Bobaru, Peridynamic model for dynamic fracture in unidirectional fiber-reinforced composites. *Comput. Methods Appl. Mech. Eng.* **217–220**, 247–261 (2012)
- B. Kilic, A. Agwai, E. Madenci, Peridynamic theory for progressive damage prediction in center-cracked composite laminates. *Compos. Struct.* **90**, 141–151 (2009)
- B. Kilic, E. Madenci, Prediction of crack paths in a quenched glass plate by using peridynamic theory. *Int. J. Fract.* **156**, 165–177 (2009)
- D.J. Littlewood, A nonlocal approach to modeling crack nucleation in AA 7075-T651, in *Proceedings of the ASME 2011 International Mechanical Engineering Congress and Exposition*, Denver, 2011. Paper No. IMECE2011-64236
- D.J. Littlewood, K. Mish, K. Pierson, Peridynamic simulation of damage evolution for structural health monitoring, in *Proceedings of the ASME 2012 International Mechanical Engineering Congress and Exposition*, Houston, 2012. Paper No. IMECE2012-86400
- E. Oterkus, E. Madenci, Peridynamics for failure prediction in composites, in *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, Honolulu, 2012
- E. Oterkus, E. Madenci, O. Weckner, S. Silling, P. Bogert, A. Tessler, Combined finite element and peridynamic analyses for predicting failure in a stiffened composite curved panel with a central slot. *Compos. Struct.* **94**, 839–850 (2012)
- M.L. Parks, R.B. Lehoucq, S.J. Plimpton, S.A. Silling, Implementing peridynamics within a molecular dynamics code. *Comput. Phys. Commun.* **179**, 777–783 (2008)
- M.L. Parks, P. Seleson, S.J. Plimpton, S.A. Silling, R.B. Lehoucq, Peridynamics with LAMMPS: a user guide v0.3 beta, Report SAND2011-8523, Sandia National Laboratories, Albuquerque and Livermore, 2011
- M.L. Parks, D.J. Littlewood, J.A. Mitchell, S.A. Silling, Peridigm users' guide v1.0.0, Report SAND2012-7800, Sandia National Laboratories, Albuquerque and Livermore, 2012
- P.D. Seleson, *Peridynamic Multiscale Models for the Mechanics of Materials: Constitutive Relations, Upscaling from Atomistic Systems, and Interface Problems*, PhD thesis, Florida State University, 2010. Electronic Theses, Treatises and Dissertations. Paper 273
- P. Seleson, Improved one-point quadrature algorithms for two-dimensional peridynamic models based on analytical calculations. *Comput. Methods Appl. Mech. Eng.* **282**, 184–217 (2014)
- P. Seleson, D.J. Littlewood, Convergence studies in meshfree peridynamic simulations. *Comput. Math. Appl.* **71**, 2432–2448 (2016)
- P. Seleson, M.L. Parks, On the role of the influence function in the peridynamic theory. *Int. J. Multiscale Comput. Eng.* **9**, 689–706 (2011)
- S.A. Silling, Linearized theory of peridynamic states. *J. Elast.* **99**, 85–111 (2010)
- S.A. Silling, E. Askari, A meshfree method based on the peridynamic model of solid mechanics. *Comput. Struct.* **83**, 1526–1535 (2005)

- S.A. Silling, M. Epton, O. Weckner, J. Xu, E. Askari, Peridynamic states and constitutive modeling. *J. Elast.* **88**, 151–184 (2007)
- S.A. Silling, O. Weckner, E. Askari, F. Bobaru, Crack nucleation in a peridynamic solid. *Int. J. Fract.* **162**, 219–227 (2010)
- M.R. Tupek, J.J. Rimoli, R. Radovitzky, An approach for incorporating classical continuum damage models in state-based peridynamics. *Comput. Methods Appl. Mech. Eng.* **263**, 20–26 (2013)
- O. Weckner, R. Abeyaratne, The effect of long-range forces on the dynamics of a bar. *J. Mech. Phys. Solids* **53**, 705–728 (2005)
- O. Weckner, G. Brunk, M.A. Epton, S.A. Silling, E. Askari, Green's functions in non-local three-dimensional linear elasticity. *Proc. R. Soc. A* **465**, 3463–3487 (2009)
- J. Xu, A. Askari, O. Weckner, S. Silling, Peridynamic analysis of impact damage in composite laminates. *J. Aerosp. Eng. Spec. Issue Impact Mech. Compos. Mater. Aerosp. Appl.* **21**, 187–194 (2008)



Well-Posed Nonlinear Nonlocal Fracture Models Associated with Double-Well Potentials

43

Prashant K. Jha and Robert Lipton

Contents

Introduction	1418
Problem Formulation with Bond-Based Nonlinear Potentials	1419
Nonlocal Potential	1420
Weak Formulation	1422
Existence of Solutions in Hölder Space	1423
Lipschitz Continuity in the Hölder Norm and Existence of a Hölder Continuous Solution ..	1425
Proof of Proposition 1	1425
Existence of Solution in Hölder Space	1433
Existence of Solutions in the Sobolev Space H^2	1436
Lipschitz Continuity in the H^2 Norm and Existence of an H^2 Solution	1437
Peridynamic Force	1438
Local and Global Existence of Solution in $H^2 \cap L^\infty$ Space	1450
Conclusions: Convergence of Regular Solutions in the Limit of Vanishing Horizon	1453
References	1456

Abstract

In this chapter, we consider a generic class of bond-based nonlocal nonlinear potentials and formulate the evolution over suitable function spaces. The peridynamic potential considered in this work is a differentiable version of the original bond-based model introduced in Silling (J Mech Phys Solids 48(1):175–

P. K. Jha

Department of Mathematics, Louisiana State University, Baton Rouge, LA, USA

e-mail: prashant.j16@gmail.com; jha@math.lsu.edu

R. Lipton (✉)

Department of Mathematics and Center for Computation and Technology, Louisiana State University, Baton Rouge, LA, USA

e-mail: lipton@lsu.edu

209, 2000). The potential associated with the model has two wells where one well corresponds to linear elastic behavior and the other corresponds to brittle fracture (see Lipton (J Elast 117(1):21–50, 2014; 124(2):143–191, 2016)). The parameters in the potential can be directly related to the elastic tensor and fracture toughness. In this chapter we show that well-posed formulations of the model can be developed over different function spaces. Here we will consider formulations posed over Hölder spaces and Sobolev spaces. The motivation for the Hölder space formulation is to show a priori convergence for the discrete finite difference method. The motivation for the Sobolev formulation is to show a priori convergence for the finite element method. In the following chapter we will show that the discrete approximations converge to well-posed evolutions. The associated convergence rates are given explicitly in terms of time step and the size of the spatial mesh.

Keywords

Peridynamic modeling · Numerical analysis · Finite difference approximation · Finite element approximation · Stability · Convergence

Introduction

The peridynamic formulation, introduced in Silling (2000), is a nonlocal model for crack propagation in solids. The basic idea is to redefine the strain in terms of the difference quotients of the displacement field and allow for nonlocal forces acting over some finite horizon. This generalized notion of strain allows for the participation of larger class of deformations in the dynamics. The modeling introduces a natural length scale given by the size of the horizon. The force at any given material point is computed by considering the deformation of all neighboring material points contained within the horizon. For linear peridynamic formulations (Silling and Lehoucq 2008; Emmrich et al. 2013; Mengesha and Du 2014; Jha and Lipton 2017c), it is shown that as the nonlocal length scale goes to zero, the peridynamic model collapses to the elastic equilibrium and elastodynamics models. For the nonlinear model introduced in Silling (2000), one may consider a smooth version to find that the energy of the evolution recovers the energy of Linear Elastic Fracture Mechanics as the nonlocal length scale goes to zero (Lipton 2014, 2016). One of the important points of this model is the fact that as the size of the horizon goes to zero, i.e., when we tend to the local limit, the model behaves as if it is an elastic model away from the crack set. Therefore, in the limit, the model not only converges to linear elasticity in regions with small deformation but also has finite Griffith fracture energy associated with a sharp fracture set. The nonlinear potential can be calibrated so that it gives the same fracture toughness as the Linear Elastic Fracture Mechanics model. Further, the slope of the nonlinear force for small strain is specified precisely by the elastic constant of the material. These results are summarized in Lipton (2016) and Jha and Lipton (2017a). On the other hand to use this model for numerical simulation, we take advantage the regularization

given by the nonlocal formulation of the problem. With this in mind we show existence of solutions in more regular spaces for fixed but small horizon and develop a theory for the numerical simulation of fracture problems. In this chapter we present the foundations for the theory and exhibit initial data and boundary conditions so that solutions exist in the Hölder space $C^{0,\gamma}$, $\gamma \in (0, 1]$ and the Sobolev space H^2 (see Jha and Lipton 2017a,b; Diehl et al. 2016). A numerical implementation scheme using the finite difference model is proposed and demonstrated in Lipton et al. (2016). In the following chapters, we show a priori convergence for the finite difference method and finite element method. These results are reported in Jha and Lipton (2017a,b). We show that these discrete approximations converge to the well-posed evolutions described in this chapter. The associated convergence rates are given explicitly in terms of time step and the size of the spatial mesh.

In this chapter we begin by describing bond-based peridynamics and the double-well potential model. Here the nonlocal forces acting between points are given by the derivatives of the potential with respect to the strain (see section “[Problem Formulation with Bond-Based Nonlinear Potentials](#)”). The existence of a peridynamic evolution taking values in the space of Hölder continuous functions is presented in section “[Existence of Solutions in Hölder Space](#).” The proof uses the Hölder continuity of the nonlocal force with respect to the Hölder norm (see section “[Lipschitz Continuity in the Hölder Norm and Existence of a Hölder Continuous Solution](#)”). We then show the existence of a peridynamic evolution in the set of essentially bounded functions taking values in the Sobolev space H^2 , the space of functions with function values, and derivatives of order one and two that are square integrable (section “[Existence of Solutions in the Sobolev Space \$H^2\$](#) ”). As before the proof uses the Hölder continuity of the nonlocal force, but now with respect to a norm that is the sum of the H^2 norm and the L^∞ norm, see section “[Lipschitz Continuity in the \$H^2\$ Norm and Existence of an \$H^2\$ Solution](#).” We conclude the chapter observing that the well-posed evolutions over these regular spaces converge to sharp fracture evolutions posed over spaces of functions with jumps (section “[Conclusions: Convergence of Regular Solutions in the Limit of Vanishing Horizon](#)”).

Problem Formulation with Bond-Based Nonlinear Potentials

Let $D \subset \mathbb{R}^d$, for $d = 2, 3$ be the material domain with characteristic length scale of unity. Every material point $x \in D$ interacts nonlocally with all other material points inside a horizon of length $\epsilon \in (0, 1)$. Let $H_\epsilon(x)$ be the ball of radius ϵ centered at x containing all points y that interact with x . After deformation the material point x assumes position $z = x + u(x)$. In this treatment we assume infinitesimal displacements $u(x)$ so the deformed configuration is the same as the reference configuration and the linearized strain is given by

$$S = S(y, x; u) = \frac{u(y) - u(x)}{|y - x|} \cdot \frac{y - x}{|y - x|}.$$

We let t denote time and the displacement field $u(t, x)$ evolves according to Newton's second law

$$\rho \partial_{tt}^2 u(t, x) = -\nabla PD^\epsilon(u(t))(x) + b(t, x) \tag{1}$$

for all $x \in D$. Here the body force applied to the domain D can evolve with time and is denoted by $b(t, x)$. Without loss of generality, we will assume $\rho = 1$. The peridynamic force denoted by $-\nabla PD^\epsilon(u)(x)$ is given by summing up all forces acting on x

$$-\nabla PD^\epsilon(u)(x) = \frac{2}{\epsilon^d \omega_d} \int_{H_\epsilon(x)} \partial_S W^\epsilon(S, y - x) \frac{y - x}{|y - x|} dy,$$

where $\partial_S W^\epsilon$ is the force exerted on x by y and is given by the derivative of the nonlocal two-point potential $W^\epsilon(S, y - x)$ with respect to the strain and ω_d is volume of unit ball in dimension d .

Let ∂D be the boundary of material domain D . The Dirichlet boundary condition on u is

$$u(t, x) = 0 \quad \forall x \in \partial D, \forall t \in [0, T] \tag{2}$$

and initial condition is

$$u(0, x) = u_0(x) \quad \text{and} \quad \partial_t u(0, x) = v_0(x). \tag{3}$$

The initial data and solution $u(t, x)$ are extended by 0 outside D .

The total energy $\mathcal{E}^\epsilon(u)(t)$ is given by the sum of kinetic and potential energy given by

$$\mathcal{E}^\epsilon(u)(t) = \frac{1}{2} \|\dot{u}(t)\|_{L^2}^2 + PD^\epsilon(u(t)), \tag{4}$$

where potential energy PD^ϵ is given by

$$PD^\epsilon(u) = \frac{1}{2} \int_D \left[\frac{1}{\epsilon^d \omega_d} \int_{H_\epsilon(x)} W^\epsilon(S(u), y - x) dy \right] dx.$$

Nonlocal Potential

We consider potentials W^ϵ of the form

$$W^\epsilon(S, y - x) = \omega(x)\omega(y) \frac{J^\epsilon(|y - x|)}{\epsilon} f(|y - x|S^2), \tag{5}$$

where $f : \mathbb{R}^+ \rightarrow \mathbb{R}$ is assumed to be positive, smooth, and concave with the following properties

$$\lim_{r \rightarrow 0^+} \frac{f(r)}{r} = f'(0), \quad \lim_{r \rightarrow \infty} f(r) = f_\infty < \infty. \tag{6}$$

The peridynamic force $-\nabla PD^\epsilon$ is written as

$$\begin{aligned} &-\nabla PD^\epsilon(u)(x) \\ &= \frac{4}{\epsilon^{d+1}\omega_d} \int_{H_\epsilon(x)} \omega(x)\omega(y)J^\epsilon(|y-x|)f'(|y-x|S(u)^2)S(u)e_{y-x}dy, \end{aligned} \tag{7}$$

where we write $S(u) = S(y, x; u)$ and $e_{y-x} = \frac{y-x}{|y-x|}$.

The function $J^\epsilon(|y-x|)$ models the influence of separation between points y and x . Here $J^\epsilon(|y-x|) = J(|y-x|/\epsilon)$, and we define J to be zero outside the ball $\{\xi : |\xi| < 1\} = H_1(0)$ and $0 \leq J(|\xi|) \leq M$ for all $\xi \in H_1(0)$.

The boundary function $\omega(x)$ is nonnegative and takes the value 1 for points x inside D of distance ϵ away from the boundary ∂D . Inside the boundary layer of width ϵ , the function $\omega(x)$ smoothly decreases from 1 to 0 taking the value 0 on ∂D .

The potential described in Eq. 5 gives the convex-concave dependence (see Fig. 1) of $W(S, y-x)$ on the strain S for fixed $y-x$. Here the potential has a well at zero strain and has a second well at infinite strain given by the horizontal asymptote. Initially the deformation is elastic for small strains and then softens as the strain becomes large; this is illustrated in Fig. 2. The critical strain where the force between x and y begins to soften is given by $S_c(y, x) := \bar{r}/\sqrt{|y-x|}$, and the force decreases monotonically for

$$|S(y, x; u)| > S_c. \tag{8}$$

Here \bar{r} is the inflection point of $r \mapsto f(r^2)$ and is the root of the following equation:

$$f'(r^2) + 2r^2 f''(r^2) = 0. \tag{9}$$

In (Theorem 5.2, Lipton 2016), it is shown that in the limit $\epsilon \rightarrow 0$, the peridynamic solution has bounded linear elastic fracture energy, provided the initial data (u_0, v_0) has bounded linear elastic fracture energy and u_0 is bounded. The elastic constants (Lamé constant λ and μ) and energy release rate of the limiting energy are given by

$$\lambda = \mu = C_d f'(0)M_d, \quad \mathcal{G}_c = \frac{2\omega_{d-1}}{\omega_d} f_\infty M_d$$

Fig. 1 Two-point potential $W^\epsilon(S, y - x)$ as a function of strain S for fixed $y - x$

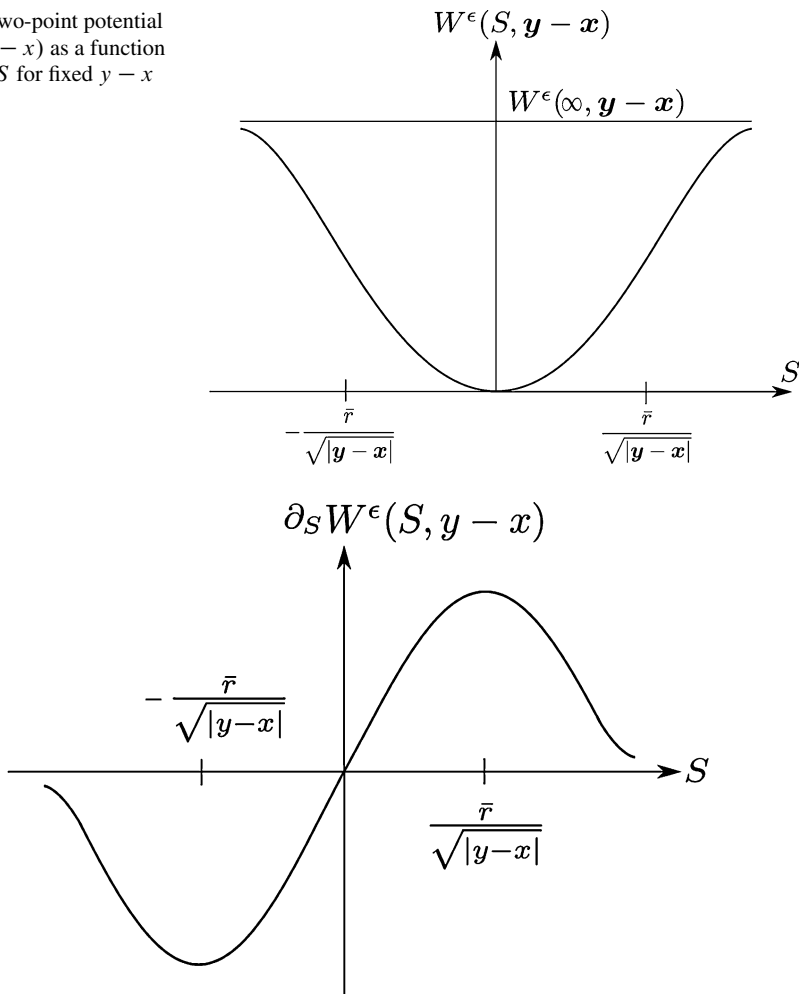


Fig. 2 Nonlocal force $\partial_S W^\epsilon(S, y - x)$ as a function of strain S for fixed $y - x$. Second derivative of $W^\epsilon(S, y - x)$ is zero at $\pm \bar{r} / \sqrt{|y - x|}$

where $M_d = \int_0^1 J(r)r^d dr$ and $f_\infty = \lim_{r \rightarrow \infty} f(r)$. $C_d = 2/3, 1/4, 1/5$ for $d = 1, 2, 3$, respectively, and $\omega_n = 1, 2, \pi, 4\pi/3$ for $n = 0, 1, 2, 3$. Therefore, $f'(0)$ and f_∞ are determined by the Lamé constant λ and fracture toughness \mathcal{G}_c .

Weak Formulation

We now give the weak formulation of the evolution. Multiplying Eq. 1 by a smooth test function \tilde{u} with $\tilde{u} = 0$ on ∂D , we get

$$(\ddot{u}(t), \tilde{u}) = (-\nabla PD^\epsilon(u(t)), \tilde{u}) + (b(t), \tilde{u}).$$

We denote L^2 dot product of u, v as (u, v) . An integration by parts easily shows for all smooth u, v taking zero boundary values that

$$(-\nabla PD^\epsilon(u), v) = -a^\epsilon(u, v),$$

where

$$\begin{aligned} a^\epsilon(u, v) &= \frac{2}{\epsilon^{d+1}\omega_d} \int_D \int_{H_\epsilon(x)} \omega(x)\omega(y)J^\epsilon(|y-x|) \\ &\quad f'(|y-x|S(u)^2)|y-x|S(u)S(v)dydx. \end{aligned} \tag{10}$$

Finally, the weak form of the evolution in terms of operator a^ϵ becomes

$$(\ddot{u}(t), \tilde{u}) + a^\epsilon(u(t), \tilde{u}) = (b(t), \tilde{u}). \tag{11}$$

Using definition of a^ϵ in Eq. 10, one easily sees that

$$\frac{d}{dt}\mathcal{E}^\epsilon(u)(t) = (\ddot{u}(t), \dot{u}(t)) + a^\epsilon(u(t), \dot{u}(t)). \tag{12}$$

In the sequel the notation $\|\cdot\|$ denotes the L^2 norm on D , and $\|\cdot\|_\infty$ is used for the L^∞ norm on D and $\|\cdot\|_2$ for Sobolev H^2 norm on D .

Existence of Solutions in Hölder Space

In this section, we establish the existence of solutions in Hölder space. Here we follow the approach developed in Jha and Lipton (2017a). Let $C^{0,\gamma}(D; \mathbb{R}^d)$ be the Hölder space with exponent $\gamma \in (0, 1]$. The closure of continuous functions with compact support on D in the supremum norm is denoted by $C_0(D)$. We identify functions in $C_0(D)$ with their unique continuous extensions to \bar{D} . It is easily seen that functions belonging to this space take the value zero on the boundary of D (see, e.g., Driver 2003). We introduce $C_0^{0,\gamma}(D) = C^{0,\gamma}(D) \cap C_0(D)$. Here we extend all functions in $C_0^{0,\gamma}(D)$ by zero outside D . The norm of $\mathbf{u} \in C_0^{0,\gamma}(D; \mathbb{R}^d)$ is taken to be

$$\|\mathbf{u}\|_{C^{0,\gamma}(D; \mathbb{R}^d)} := \sup_{\mathbf{x} \in D} |\mathbf{u}(\mathbf{x})| + [\mathbf{u}]_{C^{0,\gamma}(D; \mathbb{R}^d)},$$

where $[\mathbf{u}]_{C^{0,\gamma}(D;\mathbb{R}^d)}$ is the Hölder semi norm and given by

$$[\mathbf{u}]_{C^{0,\gamma}(D;\mathbb{R}^d)} := \sup_{\substack{x \neq y, \\ x, y \in D}} \frac{|\mathbf{u}(x) - \mathbf{u}(y)|}{|x - y|^\gamma},$$

and $C_0^{0,\gamma}(D; \mathbb{R}^d)$ is a Banach space with this norm. Here we make the hypothesis that the domain function ω belongs to $C_0^{0,\gamma}(D; \mathbb{R}^d)$.

We write the evolution Eq. 1 as an equivalent first-order system with $y_1(t) = \mathbf{u}(t)$ and $y_2(t) = \mathbf{v}(t)$ with $\mathbf{v}(t) = \partial_t \mathbf{u}(t)$. Let $y = (y_1, y_2)^T$ where $y_1, y_2 \in C_0^{0,\gamma}(D; \mathbb{R}^d)$ and let $F^\epsilon(y, t) = (F_1^\epsilon(y, t), F_2^\epsilon(y, t))^T$ such that

$$F_1^\epsilon(y, t) := y_2 \tag{13}$$

$$F_2^\epsilon(y, t) := -\nabla P D^\epsilon(y_1) + \mathbf{b}(t). \tag{14}$$

The initial boundary value associated with the evolution Eq. 1 is equivalent to the initial boundary value problem for the first-order system given by

$$\frac{d}{dt}y = F^\epsilon(y, t), \tag{15}$$

with initial condition given by $y(0) = (\mathbf{u}_0, \mathbf{v}_0)^T \in C_0^{0,\gamma}(D; \mathbb{R}^d) \times C_0^{0,\gamma}(D; \mathbb{R}^d)$.

The function $F^\epsilon(y, t)$ satisfies the Lipschitz continuity given by the following theorem.

Proposition 1 (Lipschitz continuity and bound). *Let $X = C_0^{0,\gamma}(D; \mathbb{R}^d) \times C_0^{0,\gamma}(D; \mathbb{R}^d)$. The function $F^\epsilon(y, t) = (F_1^\epsilon, F_2^\epsilon)^T$, as defined in Eqs. 13 and 14, is Lipschitz continuous in any bounded subset of X . We have, for any $y, z \in X$ and $t > 0$,*

$$\begin{aligned} & \|F^\epsilon(y, t) - F^\epsilon(z, t)\|_X \\ & \leq \frac{(L_1 + L_2 (\|\omega\|_{C^{0,\gamma}(D)} + \|y\|_X + \|z\|_X))}{\epsilon^{2+\alpha(\gamma)}} \|y - z\|_X \end{aligned} \tag{16}$$

where L_1, L_2 are independent of \mathbf{u}, \mathbf{v} and depend on peridynamic potential function f and influence function J and the exponent $\alpha(\gamma)$ is given by

$$\alpha(\gamma) = \begin{cases} 0 & \text{if } \gamma \geq 1/2 \\ 1/2 - \gamma & \text{if } \gamma < 1/2. \end{cases}$$

Furthermore for any $y \in X$ and any $t \in [0, T]$, we have the bound

$$\|F^\epsilon(y, t)\|_X \leq \frac{L_3}{\epsilon^{2+\alpha(\gamma)}} (1 + \|\omega\|_{C^{0,\gamma}(D)} + \|y\|_X) + b \tag{17}$$

where $b = \sup_t \|\mathbf{b}(t)\|_{C^{0,\gamma}(D;\mathbb{R}^d)}$ and L_3 is independent of y .

We easily see that on choosing $z = 0$ in Eq. 16 that $-\nabla PD^\epsilon(\mathbf{u})(\mathbf{x})$ is in $C^{0,\gamma}(D;\mathbb{R}^3)$ provided that \mathbf{u} belongs to $C^{0,\gamma}(D;\mathbb{R}^3)$. Since $-\nabla PD^\epsilon(\mathbf{u})(\mathbf{x})$ takes the value 0 on ∂D , we conclude that $-\nabla PD^\epsilon(\mathbf{u})(\mathbf{x})$ belongs to $C_0^{0,\gamma}(D;\mathbb{R}^3)$.

The following theorem gives the existence and uniqueness of solution in any given time domain $I_0 = (-T, T)$.

Theorem 1 (Existence and uniqueness of Hölder solutions of cohesive dynamics over finite time intervals). *For any initial condition $x_0 \in X = C_0^{0,\gamma}(D;\mathbb{R}^d) \times C_0^{0,\gamma}(D;\mathbb{R}^d)$, time interval $I_0 = (-T, T)$, and right-hand side $\mathbf{b}(t)$ continuous in time for $t \in I_0$ such that $\mathbf{b}(t)$ satisfies $\sup_{t \in I_0} \|\mathbf{b}(t)\|_{C^{0,\gamma}} < \infty$, there is a unique solution $y(t) \in C^1(I_0; X)$ of*

$$y(t) = x_0 + \int_0^t F^\epsilon(y(\tau), \tau) d\tau,$$

or equivalently

$$y'(t) = F^\epsilon(y(t), t), \text{ with } y(0) = x_0,$$

where $y(t)$ and $y'(t)$ are Lipschitz continuous in time for $t \in I_0$.

The proof of this theorem is given in the following section.

Lipschitz Continuity in the Hölder Norm and Existence of a Hölder Continuous Solution

In this section, we prove Proposition 1.

Proof of Proposition 1

Let $I = [0, T]$ be the time domain and $X = C_0^{0,\gamma}(D;\mathbb{R}^d) \times C_0^{0,\gamma}(D;\mathbb{R}^d)$. Recall that $F^\epsilon(y, t) = (F_1^\epsilon(y, t), F_2^\epsilon(y, t))$, where $F_1^\epsilon(y, t) = y^2$ and $F_2^\epsilon(y, t) = -\nabla PD^\epsilon(y^1) + \mathbf{b}(t)$. Given $t \in I$ and $y = (y^1, y^2), z = (z^1, z^2) \in X$, we have

$$\begin{aligned} & \|F^\epsilon(y, t) - F^\epsilon(z, t)\|_X \\ & \leq \|y^2 - z^2\|_{C^{0,\gamma}(D;\mathbb{R}^d)} + \|-\nabla PD^\epsilon(y^1) + \nabla PD^\epsilon(z^1)\|_{C^{0,\gamma}(D;\mathbb{R}^d)}. \end{aligned} \quad (18)$$

Therefore, to prove the Eq. 16, we only need to analyze the second term in above inequality. Let $\mathbf{u}, \mathbf{v} \in C_0^{0,\gamma}(D;\mathbb{R}^d)$, then we have

$$\begin{aligned}
 & \|-\nabla PD^\epsilon(\mathbf{u}) - (-\nabla PD^\epsilon(\mathbf{v}))\|_{C^{0,\gamma}(D;\mathbb{R}^d)} \\
 &= \sup_{\mathbf{x} \in D} |-\nabla PD^\epsilon(\mathbf{u})(\mathbf{x}) - (-\nabla PD^\epsilon(\mathbf{v})(\mathbf{x}))| \\
 & \quad + \sup_{\substack{\mathbf{x} \neq \mathbf{y}, \\ \mathbf{x}, \mathbf{y} \in D}} \frac{|(-\nabla PD^\epsilon(\mathbf{u}) + \nabla PD^\epsilon(\mathbf{v}))(\mathbf{x}) - (-\nabla PD^\epsilon(\mathbf{u}) + \nabla PD^\epsilon(\mathbf{v}))(\mathbf{y})|}{|\mathbf{x} - \mathbf{y}|^\gamma}.
 \end{aligned}
 \tag{19}$$

Note that the force $-\nabla PD^\epsilon(\mathbf{u})(\mathbf{x})$ can be written as follows:

$$\begin{aligned}
 & -\nabla PD^\epsilon(\mathbf{u})(\mathbf{x}) \\
 &= \frac{4}{\epsilon^{d+1}\omega_d} \int_{H_\epsilon(\mathbf{x})} \omega(\mathbf{x})\omega(\mathbf{y})J\left(\frac{|\mathbf{y}-\mathbf{x}|}{\epsilon}\right)f'(|\mathbf{y}-\mathbf{x}|S(\mathbf{y}, \mathbf{x}; \mathbf{u})^2)S(\mathbf{y}, \mathbf{x}; \mathbf{u})\frac{\mathbf{y}-\mathbf{x}}{|\mathbf{y}-\mathbf{x}|}d\mathbf{y} \\
 &= \frac{4}{\epsilon\omega_d} \int_{H_1(\mathbf{0})} \omega(\mathbf{x})\omega(\mathbf{x}+\epsilon\xi)J(|\xi|)f'(\epsilon|\xi|S(\mathbf{x}+\epsilon\xi, \mathbf{x}; \mathbf{u})^2)S(\mathbf{x}+\epsilon\xi, \mathbf{x}; \mathbf{u})\frac{\xi}{|\xi|}d\xi.
 \end{aligned}$$

where we substituted $\partial_S W^\epsilon$ using Eq. 5. In the second step, we introduced the change in variable $\mathbf{y} = \mathbf{x} + \epsilon\xi$.

Let $F_1 : \mathbb{R} \rightarrow \mathbb{R}$ be defined as $F_1(S) = f(S^2)$. Then $F_1'(S) = f'(S^2)2S$. Using the definition of F_1 , we have

$$2Sf'(\epsilon|\xi|S^2) = \frac{F_1'(\sqrt{\epsilon|\xi|}S)}{\sqrt{\epsilon|\xi|}}.$$

Because f is assumed to be positive, smooth, and concave and is bounded far away, we have the following bound on derivatives of F_1

$$\sup_r |F_1'(r)| = F_1'(\bar{r}) =: C_1 \tag{20}$$

$$\sup_r |F_1''(r)| = \max\{F_1''(0), F_1''(\hat{u})\} =: C_2 \tag{21}$$

$$\sup_r |F_1'''(r)| = \max\{F_1'''(\tilde{u}_2), F_1'''(\tilde{u}_2)\} =: C_3. \tag{22}$$

where \bar{r} is the inflection point of $f(r^2)$, i.e., $F_1''(\bar{r}) = 0$. $\{0, \hat{u}\}$ are the maxima of $F_1''(r)$. $\{\tilde{u}, \tilde{u}\}$ are the maxima of $F_1'''(r)$. By chain rule and by considering the assumption on f , we can show that $\bar{r}, \hat{u}, \tilde{u}_2, \tilde{u}_2$ exists and the C_1, C_2, C_3 are bounded. Figures 3, 4, and 5 show the generic graphs of $F_1'(r), F_1''(r),$ and $F_1'''(r)$, respectively.

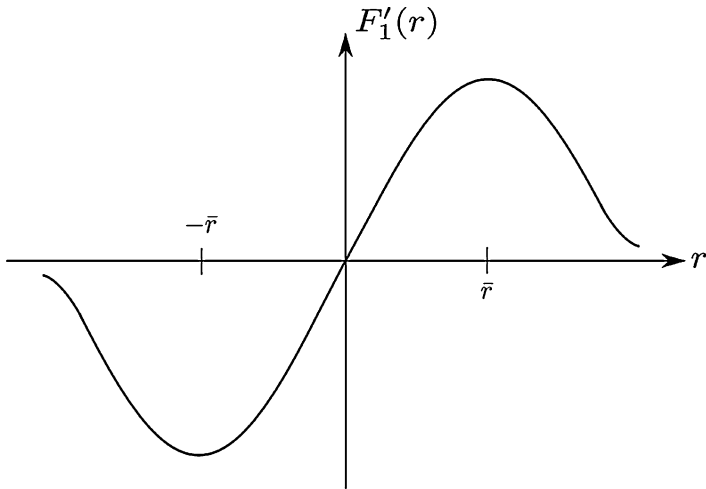


Fig. 3 Generic plot of $F'_1(r)$. $|F'_1(r)|$ is bounded by $|F'_1(\bar{r})|$

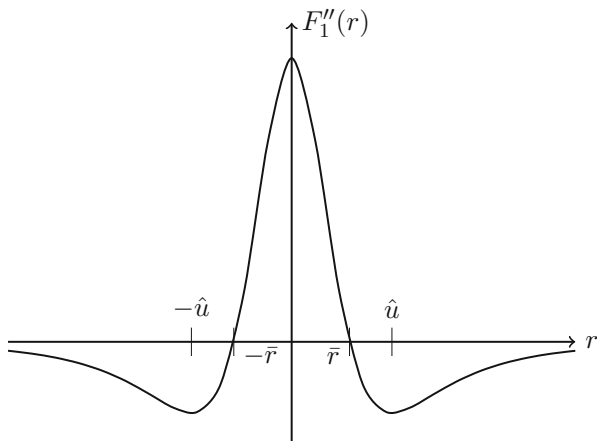


Fig. 4 Generic plot of $F''_1(r)$. At $\pm\bar{r}$, $F''_1(r) = 0$. At $\pm\hat{u}$, $F'''_1(r) = 0$

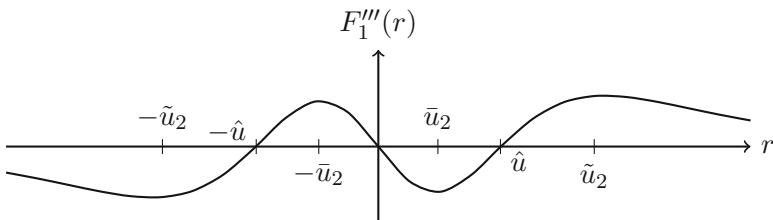


Fig. 5 Generic plot of $F'''_1(r)$. At $\pm\bar{u}_2$ and $\pm\tilde{u}_2$, $F'''_1(r) = 0$

The nonlocal force $-\nabla PD^\epsilon$ can be written as

$$\begin{aligned}
 & -\nabla PD^\epsilon(\mathbf{u})(\mathbf{x}) \\
 &= \frac{2}{\epsilon\omega_d} \int_{H_1(\mathbf{0})} \omega(\mathbf{x})\omega(\mathbf{x} + \epsilon\xi)J(|\xi|)F'_1(\sqrt{\epsilon|\xi|}S(\mathbf{x} + \epsilon\xi, \mathbf{x}; \mathbf{u})) \frac{1}{\sqrt{\epsilon|\xi|}} \frac{\xi}{|\xi|} d\xi.
 \end{aligned} \tag{23}$$

To simplify the calculations, we use following notation:

$$\begin{aligned}
 \bar{\mathbf{u}}(\mathbf{x}) &:= \mathbf{u}(\mathbf{x} + \epsilon\xi) - \mathbf{u}(\mathbf{x}), \\
 \bar{\mathbf{u}}(\mathbf{y}) &:= \mathbf{u}(\mathbf{y} + \epsilon\xi) - \mathbf{u}(\mathbf{y}), \\
 (\mathbf{u} - \mathbf{v})(\mathbf{x}) &:= \mathbf{u}(\mathbf{x}) - \mathbf{v}(\mathbf{x}),
 \end{aligned}$$

and $\overline{(\mathbf{u} - \mathbf{v})}(\mathbf{x})$ is defined similar to $\bar{\mathbf{u}}(\mathbf{x})$. Also, let

$$s = \epsilon|\xi|, \quad \mathbf{e} = \frac{\xi}{|\xi|}.$$

In what follows, we will come across the integral of type $\int_{H_1(\mathbf{0})} J(|\xi|) |\xi|^{-\alpha} d\xi$. Recall that $0 \leq J(|\xi|) \leq M$ for all $\xi \in H_1(\mathbf{0})$ and $J(|\xi|) = 0$ for $\xi \notin H_1(\mathbf{0})$. Therefore, let

$$\bar{J}_\alpha := \frac{1}{\omega_d} \int_{H_1(\mathbf{0})} J(|\xi|) |\xi|^{-\alpha} d\xi. \tag{24}$$

With notations above, we note that $S(\mathbf{x} + \epsilon\xi, \mathbf{x}; \mathbf{u}) = \bar{\mathbf{u}}(\mathbf{x}) \cdot \mathbf{e}/s$. $-\nabla PD^\epsilon$ can be written as

$$-\nabla PD^\epsilon(\mathbf{u})(\mathbf{x}) = \frac{2}{\epsilon\omega_d} \int_{H_1(\mathbf{0})} \omega(\mathbf{x})\omega(\mathbf{x} + \epsilon\xi)J(|\xi|)F'_1(\bar{\mathbf{u}}(\mathbf{x}) \cdot \mathbf{e}/\sqrt{s}) \frac{1}{\sqrt{s}} \mathbf{e} d\xi. \tag{25}$$

We first estimate the term $|\nabla PD^\epsilon(\mathbf{u})(\mathbf{x}) - \nabla PD^\epsilon(\mathbf{v})(\mathbf{x})|$ in Eq. 19.

$$\begin{aligned}
 & |-\nabla PD^\epsilon(\mathbf{u})(\mathbf{x}) - (-\nabla PD^\epsilon(\mathbf{v})(\mathbf{x}))| \\
 &\leq \left| \frac{2}{\epsilon\omega_d} \int_{H_1(\mathbf{0})} \omega(\mathbf{x})\omega(\mathbf{x} + \epsilon\xi)J(|\xi|) \frac{(F'_1(\bar{\mathbf{u}}(\mathbf{x}) \cdot \mathbf{e}/\sqrt{s}) - F'_1(\bar{\mathbf{v}}(\mathbf{x}) \cdot \mathbf{e}/\sqrt{s}))}{\sqrt{s}} \mathbf{e} d\xi \right| \\
 &\leq \left| \frac{2}{\epsilon\omega_d} \int_{H_1(\mathbf{0})} J(|\xi|) \frac{1}{\sqrt{s}} |F'_1(\bar{\mathbf{u}}(\mathbf{x}) \cdot \mathbf{e}/\sqrt{s}) - F'_1(\bar{\mathbf{v}}(\mathbf{x}) \cdot \mathbf{e}/\sqrt{s})| d\xi \right|
 \end{aligned}$$

$$\begin{aligned}
&\leq \sup_r |F_1''(r)| \left| \frac{2}{\epsilon \omega_d} \int_{H_1(\mathbf{0})} J(|\xi|) \frac{1}{\sqrt{s}} |\bar{\mathbf{u}}(\mathbf{x}) \cdot \mathbf{e} / \sqrt{s} - \bar{\mathbf{v}}(\mathbf{x}) \cdot \mathbf{e} / \sqrt{s}| d\xi \right| \\
&\leq \frac{2C_2}{\epsilon \omega_d} \left| \int_{H_1(\mathbf{0})} J(|\xi|) \frac{|\bar{\mathbf{u}}(\mathbf{x}) - \bar{\mathbf{v}}(\mathbf{x})|}{\epsilon |\xi|} d\xi \right|. \tag{26}
\end{aligned}$$

Here we have used the fact that $|\omega(\mathbf{x})| \leq 1$ and for a vector \mathbf{e} such that $|\mathbf{e}| = 1$, $|\mathbf{a} \cdot \mathbf{e}| \leq |\mathbf{a}|$ holds and $|\alpha \mathbf{e}| \leq |\alpha|$ holds for all $\mathbf{a} \in \mathbb{R}^d, \alpha \in \mathbb{R}$. Using the fact that $\mathbf{u}, \mathbf{v} \in C_0^{0,\gamma}(D; \mathbb{R}^d)$, we have

$$\begin{aligned}
\frac{|\bar{\mathbf{u}}(\mathbf{x}) - \bar{\mathbf{v}}(\mathbf{x})|}{s} &= \frac{|(\mathbf{u} - \mathbf{v})(\mathbf{x} + \epsilon \xi) - (\mathbf{u} - \mathbf{v})(\mathbf{x})|}{(\epsilon |\xi|)^\gamma} \frac{1}{(\epsilon |\xi|)^{1-\gamma}} \\
&\leq \|\mathbf{u} - \mathbf{v}\|_{C^{0,\gamma}(D; \mathbb{R}^d)} \frac{1}{(\epsilon |\xi|)^{1-\gamma}}.
\end{aligned}$$

Substituting the estimate given above, we get

$$|-\nabla PD^\epsilon(\mathbf{u})(\mathbf{x}) - (-\nabla PD^\epsilon(\mathbf{v})(\mathbf{x}))| \leq \frac{2C_2 \bar{J}_{1-\gamma}}{\epsilon^{2-\gamma}} \|\mathbf{u} - \mathbf{v}\|_{C^{0,\gamma}(D; \mathbb{R}^d)}, \tag{27}$$

where C_2 is given by Eq. 21 and $\bar{J}_{1-\gamma}$ is given by Eq. 24.

We now estimate the second term in Eq. 19. To simplify notation, we write $\tilde{\omega}(\mathbf{x}, \xi) = \omega(\mathbf{x})\omega(\mathbf{x} + \epsilon \xi)$ and with the help of Eq. 25, we get

$$\begin{aligned}
&\frac{1}{|\mathbf{x} - \mathbf{y}|^\gamma} |(-\nabla PD^\epsilon(\mathbf{u}) + \nabla PD^\epsilon(\mathbf{v}))(\mathbf{x}) - (-\nabla PD^\epsilon(\mathbf{u}) + \nabla PD^\epsilon(\mathbf{v}))(\mathbf{y})| \\
&= \frac{1}{|\mathbf{x} - \mathbf{y}|^\gamma} \left| \frac{2}{\epsilon \omega_d} \int_{H_1(\mathbf{0})} J(|\xi|) \frac{1}{\sqrt{s}} \times \left(\tilde{\omega}(\mathbf{x}, \xi) (F_1'(\frac{\bar{\mathbf{u}}(\mathbf{x}) \cdot \mathbf{e}}{\sqrt{s}}) - F_1'(\frac{\bar{\mathbf{v}}(\mathbf{x}) \cdot \mathbf{e}}{\sqrt{s}})) \right. \right. \\
&\quad \left. \left. - \tilde{\omega}(\mathbf{y}, \xi) (F_1'(\frac{\bar{\mathbf{u}}(\mathbf{y}) \cdot \mathbf{e}}{\sqrt{s}}) - F_1'(\frac{\bar{\mathbf{v}}(\mathbf{y}) \cdot \mathbf{e}}{\sqrt{s}})) \right) \mathbf{e} d\xi \right| \\
&\leq \frac{1}{|\mathbf{x} - \mathbf{y}|^\gamma} \left| \frac{2}{\epsilon \omega_d} \int_{H_1(\mathbf{0})} J(|\xi|) \frac{1}{\sqrt{s}} \times \right. \\
&\quad \left| \tilde{\omega}(\mathbf{x}, \xi) (F_1'(\frac{\bar{\mathbf{u}}(\mathbf{x}) \cdot \mathbf{e}}{\sqrt{s}}) - F_1'(\frac{\bar{\mathbf{v}}(\mathbf{x}) \cdot \mathbf{e}}{\sqrt{s}})) - \tilde{\omega}(\mathbf{y}, \xi) (F_1'(\frac{\bar{\mathbf{u}}(\mathbf{y}) \cdot \mathbf{e}}{\sqrt{s}}) \right. \\
&\quad \left. - F_1'(\frac{\bar{\mathbf{v}}(\mathbf{y}) \cdot \mathbf{e}}{\sqrt{s}})) \right| d\xi. \tag{28}
\end{aligned}$$

We analyze the integrand in above equation. We let H be defined by

$$H := \frac{|\tilde{\omega}(\mathbf{x}, \xi) (F_1'(\frac{\bar{\mathbf{u}}(\mathbf{x}) \cdot \mathbf{e}}{\sqrt{s}}) - F_1'(\frac{\bar{\mathbf{v}}(\mathbf{x}) \cdot \mathbf{e}}{\sqrt{s}})) - \tilde{\omega}(\mathbf{y}, \xi) (F_1'(\frac{\bar{\mathbf{u}}(\mathbf{y}) \cdot \mathbf{e}}{\sqrt{s}}) - F_1'(\frac{\bar{\mathbf{v}}(\mathbf{y}) \cdot \mathbf{e}}{\sqrt{s}}))|}{|\mathbf{x} - \mathbf{y}|^\gamma}.$$

Let $r : [0, 1] \times D \rightarrow \mathbb{R}^d$ be defined as

$$r(l, x) = \bar{v}(x) + l(\bar{u}(x) - \bar{v}(x)).$$

Note $\partial r(l, x)/\partial l = \bar{u}(x) - \bar{v}(x)$. Using $r(l, x)$, we have

$$F'_1(\bar{u}(x) \cdot e/\sqrt{s}) - F'_1(\bar{v}(x) \cdot e/\sqrt{s}) = \int_0^1 \frac{\partial F'_1(r(l, x) \cdot e/\sqrt{s})}{\partial l} dl \tag{29}$$

$$= \int_0^1 \frac{\partial F'_1(r \cdot e/\sqrt{s})}{\partial r} \Big|_{r=r(l,x)} \cdot \frac{\partial r(l, x)}{\partial l} dl. \tag{30}$$

Similarly, we have

$$F'_1(\bar{u}(y) \cdot e/\sqrt{s}) - F'_1(\bar{v}(y) \cdot e/\sqrt{s}) = \int_0^1 \frac{\partial F'_1(r \cdot e/\sqrt{s})}{\partial r} \Big|_{r=r(l,y)} \cdot \frac{\partial r(l, y)}{\partial l} dl. \tag{31}$$

Note that

$$\frac{\partial F'_1(r \cdot e/\sqrt{s})}{\partial r} \Big|_{r=r(l,y)} = F''_1(r(l, x) \cdot e/\sqrt{s}) \frac{e}{\sqrt{s}}. \tag{32}$$

Combining Eqs. 30, 31, and 32 gives

$$\begin{aligned} H &= \frac{1}{|x - y|^\gamma} \left| \int_0^1 (\tilde{\omega}(x, \xi) F''_1(r(l, x) \cdot e/\sqrt{s})(\bar{u}(x) - \bar{v}(x)) \right. \\ &\quad \left. - \tilde{\omega}(y, \xi) F''_1(r(l, y) \cdot e/\sqrt{s})(\bar{u}(y) - \bar{v}(y))) \cdot \frac{e}{\sqrt{s}} dl \right| \\ &\leq \frac{1}{|x - y|^\gamma} \frac{1}{\sqrt{s}} \left| \int_0^1 |\tilde{\omega}(x, \xi) F''_1(r(l, x) \cdot e/\sqrt{s})(\bar{u}(x) - \bar{v}(x)) \right. \\ &\quad \left. - \tilde{\omega}(y, \xi) F''_1(r(l, y) \cdot e/\sqrt{s})(\bar{u}(y) - \bar{v}(y))| dl \right|. \end{aligned}$$

Adding and subtracting $\tilde{\omega}(x, \xi) F''_1(r(l, x) \cdot e/\sqrt{s})(\bar{u}(y) - \bar{v}(y))$ and noting $0 \leq \tilde{\omega}(x, \xi) \leq 1$ give

$$\begin{aligned} H &\leq \frac{1}{|x - y|^\gamma} \frac{1}{\sqrt{s}} \left| \int_0^1 |F''_1(r(l, x) \cdot e/\sqrt{s})| |\bar{u}(x) - \bar{v}(x) - \bar{u}(y) + \bar{v}(y)| dl \right| \\ &\quad + \frac{1}{|x - y|^\gamma} \frac{1}{\sqrt{s}} \int_0^1 |(\tilde{\omega}(x, \xi) F''_1(r(l, x) \cdot e/\sqrt{s}) - \tilde{\omega}(y, \xi) F''_1(r(l, y) \cdot e/\sqrt{s}))| \\ &\quad \times |\bar{u}(y) - \bar{v}(y)| dl. \\ &=: H_1 + H_2. \end{aligned}$$

The H_1 term is estimated first. Note that $|F_1''(r)| \leq C_2$. Since $\mathbf{u}, \mathbf{v} \in C_0^{0,\gamma}(D; \mathbb{R}^d)$, it is easily seen that

$$\frac{|\bar{\mathbf{u}}(\mathbf{x}) - \bar{\mathbf{v}}(\mathbf{x}) - \bar{\mathbf{u}}(\mathbf{y}) + \bar{\mathbf{v}}(\mathbf{y})|}{|\mathbf{x} - \mathbf{y}|^\gamma} \leq 2\|\mathbf{u} - \mathbf{v}\|_{C^{0,\gamma}(D; \mathbb{R}^d)}.$$

Therefore, we have

$$H_1 \leq \frac{2C_2}{\sqrt{s}} \|\mathbf{u} - \mathbf{v}\|_{C^{0,\gamma}(D; \mathbb{R}^d)}. \tag{33}$$

We now estimate H_2 . We add and subtract $\tilde{\omega}(\mathbf{x}, \boldsymbol{\xi})F_1''(\mathbf{r}(l, \mathbf{y}) \cdot \mathbf{e} / \sqrt{s})$ in H_2 to get

$$H_2 \leq H_3 + H_4,$$

where

$$H_3 = \frac{1}{|\mathbf{x} - \mathbf{y}|^\gamma} \frac{1}{\sqrt{s}} \int_0^1 |(F_1''(\mathbf{r}(l, \mathbf{x}) \cdot \mathbf{e} / \sqrt{s}) - F_1''(\mathbf{r}(l, \mathbf{y}) \cdot \mathbf{e} / \sqrt{s}))| |\bar{\mathbf{u}}(\mathbf{y}) - \bar{\mathbf{v}}(\mathbf{y})| dl,$$

and

$$H_4 = \frac{1}{|\mathbf{x} - \mathbf{y}|^\gamma} \frac{1}{\sqrt{s}} \int_0^1 |(\tilde{\omega}(\mathbf{x}, \boldsymbol{\xi}) - \tilde{\omega}(\mathbf{y}, \boldsymbol{\xi}))F_1''(\mathbf{r}(l, \mathbf{y}) \cdot \mathbf{e} / \sqrt{s})| |\bar{\mathbf{u}}(\mathbf{y}) - \bar{\mathbf{v}}(\mathbf{y})| dl.$$

Now we estimate H_3 . Since $|F_1'''(r)| \leq C_3$ (see Eq. 22), we have

$$\begin{aligned} & \frac{1}{|\mathbf{x} - \mathbf{y}|^\gamma} |F_1''(\mathbf{r}(l, \mathbf{x}) \cdot \mathbf{e} / \sqrt{s}) - F_1''(\mathbf{r}(l, \mathbf{y}) \cdot \mathbf{e} / \sqrt{s})| \\ & \leq \frac{1}{|\mathbf{x} - \mathbf{y}|^\gamma} \sup_r |F_1'''(r)| \frac{|\mathbf{r}(l, \mathbf{x}) \cdot \mathbf{e} - \mathbf{r}(l, \mathbf{y}) \cdot \mathbf{e}|}{\sqrt{s}} \\ & \leq \frac{C_3}{\sqrt{s}} \frac{|\mathbf{r}(l, \mathbf{x}) - \mathbf{r}(l, \mathbf{y})|}{|\mathbf{x} - \mathbf{y}|^\gamma} \\ & = \frac{C_3}{\sqrt{s}} \left(\frac{|1 - l| |\bar{\mathbf{v}}(\mathbf{x}) - \bar{\mathbf{v}}(\mathbf{y})|}{|\mathbf{x} - \mathbf{y}|^\gamma} + \frac{|l| |\bar{\mathbf{u}}(\mathbf{x}) - \bar{\mathbf{u}}(\mathbf{y})|}{|\mathbf{x} - \mathbf{y}|^\gamma} \right) \\ & \leq \frac{C_3}{\sqrt{s}} \left(\frac{|\bar{\mathbf{v}}(\mathbf{x}) - \bar{\mathbf{v}}(\mathbf{y})|}{|\mathbf{x} - \mathbf{y}|^\gamma} + \frac{|\bar{\mathbf{u}}(\mathbf{x}) - \bar{\mathbf{u}}(\mathbf{y})|}{|\mathbf{x} - \mathbf{y}|^\gamma} \right). \end{aligned} \tag{34}$$

Where we have used the fact that $|1 - l| \leq 1, |l| \leq 1, \text{ as } l \in [0, 1]$. Also, note that

$$\begin{aligned} \frac{|\bar{\mathbf{u}}(\mathbf{x}) - \bar{\mathbf{u}}(\mathbf{y})|}{|\mathbf{x} - \mathbf{y}|^\gamma} &\leq 2\|\mathbf{u}\|_{C^{0,\gamma}(D;\mathbb{R}^d)} \\ \frac{|\bar{\mathbf{v}}(\mathbf{x}) - \bar{\mathbf{v}}(\mathbf{y})|}{|\mathbf{x} - \mathbf{y}|^\gamma} &\leq 2\|\mathbf{v}\|_{C^{0,\gamma}(D;\mathbb{R}^d)} \\ |\bar{\mathbf{u}}(\mathbf{y}) - \bar{\mathbf{v}}(\mathbf{y})| &\leq s^\gamma \|\mathbf{u} - \mathbf{v}\|_{C^{0,\gamma}(D;\mathbb{R}^d)}. \end{aligned}$$

We combine above estimates with Eq. 34 to get

$$\begin{aligned} H_3 &\leq \frac{1}{\sqrt{s}} \frac{C_3}{\sqrt{s}} (\|\mathbf{u}\|_{C^{0,\gamma}(D;\mathbb{R}^d)} + \|\mathbf{v}\|_{C^{0,\gamma}(D;\mathbb{R}^d)}) s^\gamma \|\mathbf{u} - \mathbf{v}\|_{C^{0,\gamma}(D;\mathbb{R}^d)} \\ &= \frac{C_3}{s^{1-\gamma}} (\|\mathbf{u}\|_{C^{0,\gamma}(D;\mathbb{R}^d)} + \|\mathbf{v}\|_{C^{0,\gamma}(D;\mathbb{R}^d)}) \|\mathbf{u} - \mathbf{v}\|_{C^{0,\gamma}(D;\mathbb{R}^d)}. \end{aligned} \tag{35}$$

Next we estimate H_4 . Here we add and subtract $\omega(\mathbf{y})\omega(\mathbf{x} + \epsilon\xi)$ to get

$$\begin{aligned} H_4 &= \frac{1}{|\mathbf{x} - \mathbf{y}|^\gamma} \frac{1}{\sqrt{s}} \int_0^1 |(\omega(\mathbf{x}, \mathbf{x} + \epsilon\xi)(\omega(\mathbf{x}) - \omega(\mathbf{y})) + \omega(\mathbf{y})(\omega(\mathbf{x} + \epsilon\xi) - \omega(\mathbf{y} + \epsilon\xi))) \\ &\quad \times |F_1''(\mathbf{r}(l, \mathbf{y}) \cdot \mathbf{e} / \sqrt{s})| |\bar{\mathbf{u}}(\mathbf{y}) - \bar{\mathbf{v}}(\mathbf{y})| dl. \end{aligned}$$

Recalling that ω belongs to $C_0^{0,\gamma}(D; \mathbb{R}^d)$ and in view of the previous estimates, a straightforward calculation gives

$$H_4 \leq \frac{4C_2}{s^{1/2-\gamma}} \|\omega\|_{C^{0,\gamma}(D;\mathbb{R}^d)} \|\mathbf{u} - \mathbf{v}\|_{C^{0,\gamma}(D;\mathbb{R}^d)}. \tag{36}$$

Combining Eqs. 33, 35, and 36 gives

$$\begin{aligned} H &\leq \left(\frac{2C_2}{\sqrt{s}} + \frac{4C_2}{s^{1/2-\gamma}} \|\omega\|_{C^{0,\gamma}(D;\mathbb{R}^d)} + \right. \\ &\quad \left. + \frac{C_3}{s^{1-\gamma}} (\|\mathbf{u}\|_{C^{0,\gamma}(D;\mathbb{R}^d)} + \|\mathbf{v}\|_{C^{0,\gamma}(D;\mathbb{R}^d)}) \right) \|\mathbf{u} - \mathbf{v}\|_{C^{0,\gamma}(D;\mathbb{R}^d)}. \end{aligned}$$

Substituting H in Eq. 28 gives

$$\begin{aligned} &\frac{1}{|\mathbf{x} - \mathbf{y}|^\gamma} |(-\nabla PD^\epsilon(\mathbf{u}) + \nabla PD^\epsilon(\mathbf{v}))(\mathbf{x}) - (-\nabla PD^\epsilon(\mathbf{u}) + \nabla PD^\epsilon(\mathbf{v}))(\mathbf{y})| \\ &\leq \left| \frac{2}{\epsilon\omega_d} \int_{H_1(\mathbf{0})} J(|\xi|) \frac{1}{\sqrt{s}} Hd\xi \right| \end{aligned}$$

$$\begin{aligned} &\leq \left(\frac{4C_2\bar{J}_1}{\epsilon^2} + \frac{4C_2\bar{J}_{1-\gamma}}{\epsilon^{2-\gamma}} \|\omega\|_{C^{0,\gamma}(D;\mathbb{R}^d)} \right. \\ &\quad \left. + \frac{2C_3\bar{J}_{3/2-\gamma}}{\epsilon^{2+1/2-\gamma}} (\|\mathbf{u}\|_{C^{0,\gamma}(D;\mathbb{R}^d)} + \|\mathbf{v}\|_{C^{0,\gamma}(D;\mathbb{R}^d)}) \right) \|\mathbf{u} - \mathbf{v}\|_{C^{0,\gamma}(D;\mathbb{R}^d)}. \end{aligned} \tag{37}$$

We combine Eqs. 19, 27, and 37 and get

$$\begin{aligned} &\|-\nabla PD^\epsilon(\mathbf{u}) - (-\nabla PD^\epsilon(\mathbf{v}))\|_{C^{0,\gamma}} \\ &\leq \left(\frac{4C_2\bar{J}_1}{\epsilon^2} + \frac{2C_2\bar{J}_{1-\gamma}}{\epsilon^{2-\gamma}}(1 + \|\omega\|_{C^{0,\gamma}}) + \frac{2C_3\bar{J}_{3/2-\gamma}}{\epsilon^{2+1/2-\gamma}} (\|\mathbf{u}\|_{C^{0,\gamma}} + \|\mathbf{v}\|_{C^{0,\gamma}}) \right) \|\mathbf{u} - \mathbf{v}\|_{C^{0,\gamma}} \\ &\leq \frac{\bar{C}_1 + \bar{C}_2\|\omega\|_{C^{0,\gamma}} + \bar{C}_3(\|\mathbf{u}\|_{C^{0,\gamma}} + \|\mathbf{v}\|_{C^{0,\gamma}})}{\epsilon^{2+\alpha(\gamma)}} \|\mathbf{u} - \mathbf{v}\|_{C^{0,\gamma}} \end{aligned} \tag{38}$$

where we introduce new constants $\bar{C}_1, \bar{C}_2, \bar{C}_3$. We let $\alpha(\gamma) = 0$, if $\gamma \geq 1/2$, and $\alpha(\gamma) = 1/2 - \gamma$, if $\gamma \leq 1/2$. One can easily verify that, for all $\gamma \in (0, 1]$ and $0 < \epsilon \leq 1$,

$$\max \left\{ \frac{1}{\epsilon^2}, \frac{1}{\epsilon^{2+1/2-\gamma}}, \frac{1}{\epsilon^{2-\gamma}} \right\} \leq \frac{1}{\epsilon^{2+\alpha(\gamma)}}$$

To complete the proof, we combine Eqs. 38 and 18 and get

$$\|F^\epsilon(y, t) - F^\epsilon(z, t)\|_X \leq \frac{L_1 + L_2(\|\omega\|_{C^{0,\gamma}} + \|y\|_X + \|z\|_X)}{\epsilon^{2+\alpha(\gamma)}} \|y - z\|_X.$$

This proves the Lipschitz continuity of $F^\epsilon(y, t)$ on any bounded subset of X . The bound on $F^\epsilon(y, t)$ (see Eq. 17) follows easily from Eq. 25. This completes the proof of Proposition 1.

Existence of Solution in Hölder Space

In this section, we prove Theorem 1. We begin by proving a local existence theorem. We then show that the local solution can be continued uniquely in time to recover Theorem 1.

The existence and uniqueness of local solutions is stated in the following theorem.

Theorem 2 (Local existence and uniqueness). *Given $X = C_0^{0,\gamma}(D; \mathbb{R}^d) \times C_0^{0,\gamma}(D; \mathbb{R}^d)$, $\mathbf{b}(t) \in C_0^{0,\gamma}(D; \mathbb{R}^d)$, and initial data $x_0 = (\mathbf{u}_0, \mathbf{v}_0) \in X$. We suppose that $\mathbf{b}(t)$ is continuous in time over some time interval $I_0 = (-T, T)$*

and satisfies $\sup_{t \in I_0} \|\mathbf{b}(t)\|_{C^{0,\gamma}(D; \mathbb{R}^d)} < \infty$. Then, there exists a time interval $I' = (-T', T') \subset I_0$ and unique solution $y = (y^1, y^2)$ such that $y \in C^1(I'; X)$ and

$$y(t) = x_0 + \int_0^t F^\epsilon(y(\tau), \tau) d\tau, \text{ for } t \in I' \tag{39}$$

or equivalently

$$y'(t) = F^\epsilon(y(t), t), \text{ with } y(0) = x_0, \text{ for } t \in I'$$

where $y(t)$ and $y'(t)$ are Lipschitz continuous in time for $t \in I' \subset I_0$.

To prove Theorem 2, we proceed as follows. We write $y(t) = (y^1(t), y^2(t))$ and $\|y\|_X = \|y^1(t)\|_{C^{0,\gamma}} + \|y^2(t)\|_{C^{0,\gamma}}$. Define the ball $B(0, R) = \{y \in X : \|y\|_X < R\}$ and choose $R > \|x_0\|_X$. Let $r = R - \|x_0\|_X$ and we consider the ball $B(x_0, r)$ defined by

$$B(x_0, r) = \{y \in X : \|y - x_0\|_X < r\} \subset B(0, R), \tag{40}$$

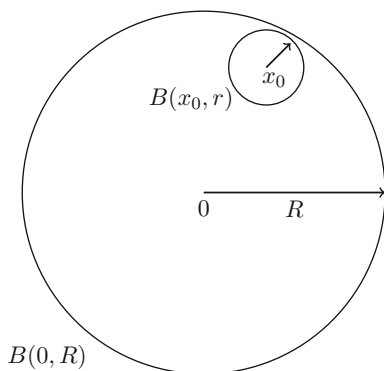
(see Fig. 6).

To recover the existence and uniqueness, we introduce the transformation

$$S_{x_0}(y)(t) = x_0 + \int_0^t F^\epsilon(y(\tau), \tau) d\tau.$$

Introduce $0 < T' < T$ and the associated set $Y(T')$ of Hölder continuous functions taking values in $B(x_0, r)$ for $I' = (-T', T') \subset I_0 = (-T, T)$. The goal is to find appropriate interval $I' = (-T', T')$ for which S_{x_0} maps into the corresponding set $Y(T')$. Writing out the transformation with $y(t) \in Y(T')$ gives

Fig. 6 Geometry



$$S_{x_0}^1(y)(t) = x_0^1 + \int_0^t y^2(\tau) d\tau \tag{41}$$

$$S_{x_0}^2(y)(t) = x_0^2 + \int_0^t (-\nabla PD^\epsilon(y^1(\tau)) + \mathbf{b}(\tau)) d\tau, \tag{42}$$

and there is a positive constant $K = C/\epsilon^{2+\alpha(y)}$ (see Eq. 17) independent of $y^1(t)$, for $-T' < t < T'$, such that estimation in Eq. 42 gives

$$\|S_{x_0}^2(y)(t) - x_0^2\|_{C^{0,\gamma}} \leq (K(1 + \frac{1}{\epsilon^\gamma} + \sup_{t \in (-T', T')} \|y^1(t)\|_{C^{0,\gamma}}) + \sup_{t \in (-T, T)} \|\mathbf{b}(t)\|_{C^{0,\gamma}})T' \tag{43}$$

and from Eq. 41

$$\|S_{x_0}^1(y)(t) - x_0^1\|_{C^{0,\gamma}} \leq \sup_{t \in (-T', T')} \|y^2(t)\|_{C^{0,\gamma}} T'. \tag{44}$$

We write $b = \sup_{t \in I_0} \|\mathbf{b}(t)\|_{C^{0,\gamma}}$ and adding Eqs. 43 and 44 gives the upper bound

$$\|S_{x_0}(y)(t) - x_0\|_X \leq (K(1 + \frac{1}{\epsilon^\gamma} + \sup_{t \in (-T', T')} \|y(t)\|_X) + b)T'. \tag{45}$$

Since $B(x_0, r) \subset B(0, R)$ (see Eq. 40), we make the choice T' so that

$$\|S_{x_0}(y)(t) - x_0\|_X \leq ((K(1 + \frac{1}{\epsilon^\gamma} + R) + b)T' < r = R - \|x_0\|_X. \tag{46}$$

For this choice we see that

$$T' < \theta(R) = \frac{R - \|x_0\|_X}{K(R + 1 + \frac{1}{\epsilon^\gamma}) + b}. \tag{47}$$

Now it is easily seen that $\theta(R)$ is increasing with $R > 0$ and

$$\lim_{R \rightarrow \infty} \theta(R) = \frac{1}{K}. \tag{48}$$

So given R and $\|x_0\|_X$, we choose T' according to

$$\frac{\theta(R)}{2} < T' < \theta(R), \tag{49}$$

and set $I' = (-T', T')$. We have found the appropriate time domain I' such that the transformation $S_{x_0}(y)(t)$ as defined in Eqs. 41 and 42 maps $Y(T')$ into itself. We now proceed using standard arguments (see, e.g., Driver 2003, Theorem 6.10)

to complete the proof of existence and uniqueness of solution for given initial data x_0 over the interval $I' = (-T', T')$.

We now prove Theorem 1. From the proof of Theorem 2 above, we see that a unique local solution exists over a time domain $(-T', T')$ with $\frac{\theta(R)}{2} < T'$. Since $\theta(R) \nearrow 1/K$ as $R \nearrow \infty$, we can fix a tolerance $\eta > 0$ so that $[(1/2K) - \eta] > 0$. Then given any initial condition with bounded Hölder norm and $b = \sup_{t \in [-T, T]} \|b(t)\|_{C^{0,\gamma}}$, we can choose R sufficiently large so that $\|x_0\|_X < R$ and $0 < (1/2K) - \eta < T'$. Thus we can always find local solutions for time intervals $(-T', T')$ for T' larger than $[(1/2K) - \eta] > 0$. Therefore we apply the local existence and uniqueness result to uniquely continue local solutions up to an arbitrary time interval $(-T, T)$.

Existence of Solutions in the Sobolev Space H^2

We start by recalling that the space $H_0^2(D; \mathbb{R}^d)$ is the closure in the H^2 norm of twice differentiable functions with compact support in D . We denote the norm in H^m by $\|\cdot\|_m$, $m = 1, 2$, and the L^∞ norm by $\|\cdot\|_\infty$. In this section, we find that solutions of peridynamic evolutions exist for almost all times in $H_0^2(D; \mathbb{R}^d) \cap L^\infty(D; \mathbb{R}^d)$. For the sake of convenience, we let W denote the $H_0^2(D; \mathbb{R}^d) \cap L^\infty(D; \mathbb{R}^d)$ space. The norm on W is defined as

$$\|u\|_W := \|u\|_2 + \|u\|_\infty. \tag{50}$$

We will assume that $u \in H_0^2(D; \mathbb{R}^d)$ is extended by zero outside D ; therefore, $u = 0, \nabla u = 0, \nabla^2 u = 0$ for $x \notin D$ and $\|u\|_{H^2(D; \mathbb{R}^d)} = \|u\|_{H^2(\mathbb{R}^d; \mathbb{R}^d)}$.

Noting the Sobolev embedding property of $u \in H_0^2(D; \mathbb{R}^d)$ (see Theorem 2.31, Demengel and Demengel 2012) given by

$$\|\nabla u\|_{L^q(D; \mathbb{R}^{d \times d})} \leq C_e \|u\|_{H_0^2(D; \mathbb{R}^d)} \tag{51}$$

for any q such that $2 \leq q \leq 6$ in case of $d = 3$ and $2 \leq q < \infty$ in case of $d = 2$. Constant C_e is independent of u .

In what follows, we will first prove the Lipschitz bound on $-\nabla PD^\epsilon(u)$, and then using Lipschitz bound, we will show the local existence of solution u in W . We write the peridynamic equation as an equivalent first-order system with $y_1(t) = u(t)$ and $y_2(t) = v(t)$ with $v(t) = \dot{u}(t)$. Let $y = (y_1, y_2)^T$ where $y_1, y_2 \in W$ and let $F^\epsilon(y, t) = (F_1^\epsilon(y, t), F_2^\epsilon(y, t))^T$ such that

$$F_1^\epsilon(y, t) := y_2, \tag{52}$$

$$F_2^\epsilon(y, t) := -\nabla PD^\epsilon(y_1) + b(t). \tag{53}$$

The initial boundary value is equivalent to the initial boundary value problem for the first-order system given by

$$\dot{y}(t) = F^\epsilon(y, t), \tag{54}$$

with initial condition given by $y(0) = (u_0, v_0)^T \in W \times W$.

Theorem 3 (Lipschitz bound on peridynamic force). *For any $u, v \in W$, we have*

$$\begin{aligned} & \| -\nabla PD^\epsilon(u) - (-\nabla PD^\epsilon(v)) \|_W \\ & \leq \frac{\bar{L}_1 + \bar{L}_2(\|u\|_W + \|v\|_W) + \bar{L}_3(\|u\|_W + \|v\|_W)^2}{\epsilon^3} \|u - v\|_W \end{aligned} \tag{55}$$

where constants $\bar{L}_1, \bar{L}_2, \bar{L}_3$ are independent of ϵ, u , and v and are defined in (96). Also, for $u \in W$, we have

$$\| -\nabla PD^\epsilon(u) \|_W \leq \frac{\bar{L}_4\|u\|_W + \bar{L}_5\|u\|_W^2}{\epsilon^{5/2}}, \tag{56}$$

where constants are independent of ϵ and u and are defined in (105).

We state the theorem which shows the existence and uniqueness of solution in any given finite time interval $I_0 = (-T, T)$.

Theorem 4 (Existence and uniqueness of solutions over finite time intervals).

For any initial condition $x_0 \in X = W \times W$, time interval $I_0 = (-T, T)$, and right-hand side $b(t)$ continuous in time for $t \in I_0$ such that $b(t)$ satisfies $\sup_{t \in I_0} \|b(t)\|_W < \infty$, there is a unique solution $y(t) \in C^1(I_0; X)$ of

$$y(t) = x_0 + \int_0^t F^\epsilon(y(\tau), \tau) d\tau,$$

or equivalently

$$y'(t) = F^\epsilon(y(t), t), \text{ with } y(0) = x_0,$$

where $y(t)$ and $y'(t)$ are Lipschitz continuous in time for $t \in I_0$.

The proof of the Lipschitz continuity and existence is established in the following section.

Lipschitz Continuity in the H^2 Norm and Existence of an H^2 Solution

In this section we prove Theorems 3 and 4. To simplify the presentation, we denote the peridynamic force $-\nabla PD^\epsilon(u)$ by simply $P(u)$. Recall that we denote $H_0^2(D; \mathbb{R}^d) \cap L^\infty(D; \mathbb{R}^d)$ by W and

$$\|u\|_W = \|u\| + \|\nabla u\| + \|\nabla^2 u\| + \|u\|_\infty.$$

We need to analyze $\|P(u) - P(v)\|_W$.

We use the following short notations:

$$s_\xi = \epsilon|\xi|, \quad e_\xi = \frac{\xi}{|\xi|}, \quad \bar{J}_\alpha = \frac{1}{\omega_d} \int_{H_1(0)} J(|\xi|) \frac{1}{|\xi|^\alpha} d\xi,$$

$$S_\xi(u) = \frac{u(x + \epsilon\xi) - u(x)}{s_\xi} \cdot e_\xi,$$

$$S_\xi(\nabla u) = \nabla S_\xi(u) = \frac{\nabla u^T(x + \epsilon\xi) - \nabla u^T(x)}{s_\xi} e_\xi,$$

$$S_\xi(\nabla^2 u) = \nabla S_\xi(\nabla u) = \nabla \left[\frac{\nabla u^T(x + \epsilon\xi) - \nabla u^T(x)}{s_\xi} e_\xi \right].$$

In indicial notation, we have

$$S_\xi(\nabla u)_i = \frac{u_{k,i}(x + \epsilon\xi) - u_{k,i}(x)}{s_\xi} (e_\xi)_k,$$

$$S_\xi(\nabla^2 u)_{ij} = \left[\frac{u_{k,i}(x + \epsilon\xi) - u_{k,i}(x)}{s_\xi} (e_\xi)_k \right]_{,j} = \frac{u_{k,ij}(x + \epsilon\xi) - u_{k,ij}(x)}{s_\xi} (e_\xi)_k \tag{57}$$

and

$$[e_\xi \otimes S_\xi(\nabla^2 u)]_{ijk} = (e_\xi)_i S_\xi(\nabla^2 u)_{jk}, \tag{58}$$

where $u_{i,j} = (\nabla u)_{ij}$, $u_{k,ij} = (\nabla^2 u)_{kij}$, and $(e_\xi)_k = \xi_k/|\xi|$.

Peridynamic Force

Let $F_1(r) := f(r^2)$ where f is described in section “[Problem Formulation with Bond-Based Nonlinear Potentials](#).” We have $F'_1(r) = f'(r^2)2r$. Thus, $2Sf'(\epsilon|\xi|S^2) = F'_1(\sqrt{\epsilon|\xi|}S)/\sqrt{\epsilon|\xi|}$. We define the following constants related to nonlinear potential

$$C_1 := \sup_r |F'_1(r)|, \quad C_2 := \sup_r |F''_1(r)|, \quad C_3 := \sup_r |F'''_1(r)|, \quad C_4 := \sup_r |F''''_1(r)|.$$

The potential function f as chosen here satisfies $C_1, C_2, C_3, C_4 < \infty$. Let

$$\bar{\omega}_\xi(x) = \omega(x)\omega(x + \epsilon\xi), \tag{59}$$

and we choose ω such that

$$|\nabla \bar{\omega}_\xi| \leq C_{\omega_1} < \infty \quad \text{and} \quad |\nabla^2 \bar{\omega}_\xi| \leq C_{\omega_2} < \infty. \tag{60}$$

With notations described so far, we write peridynamic force $P(u)$ as

$$P(u)(x) = \frac{2}{\epsilon \omega_d} \int_{H_1(0)} \bar{\omega}_\xi(x) J(|\xi|) \frac{F'_1(\sqrt{s_\xi} S_\xi(u))}{\sqrt{s_\xi}} e_\xi d\xi. \tag{61}$$

The gradient of $P(u)(x)$ is given by

$$\begin{aligned} \nabla P(u)(x) &= \frac{2}{\epsilon \omega_d} \int_{H_1(0)} \bar{\omega}_\xi(x) J(|\xi|) F''_1(\sqrt{s_\xi} S_\xi(u)) e_\xi \otimes \nabla S_\xi(u) d\xi \\ &\quad + \frac{2}{\epsilon \omega_d} \int_{H_1(0)} J(|\xi|) \frac{F'_1(\sqrt{s_\xi} S_\xi(u))}{\sqrt{s_\xi}} e_\xi \otimes \nabla \bar{\omega}_\xi(x) d\xi \\ &= g_1(u)(x) + g_2(u)(x), \end{aligned} \tag{62}$$

where we denote first and second term as $g_1(u)(x)$ and $g_2(u)(x)$, respectively. We also have

$$\begin{aligned} \nabla^2 P(u)(x) &= \frac{2}{\epsilon \omega_d} \int_{H_1(0)} \bar{\omega}_\xi(x) J(|\xi|) F''_1(\sqrt{s_\xi} S_\xi(u)) e_\xi \otimes S_\xi(\nabla^2 u) d\xi \\ &\quad + \frac{2}{\epsilon \omega_d} \int_{H_1(0)} \bar{\omega}_\xi(x) J(|\xi|) \sqrt{s_\xi} F'''_1(\sqrt{s_\xi} S_\xi(u)) e_\xi \otimes S_\xi(\nabla u) \otimes S_\xi(\nabla u) d\xi \\ &\quad + \frac{2}{\epsilon \omega_d} \int_{H_1(0)} J(|\xi|) F''_1(\sqrt{s_\xi} S_\xi(u)) e_\xi \otimes S_\xi(\nabla u) \otimes \nabla \bar{\omega}_\xi(x) d\xi \\ &\quad + \frac{2}{\epsilon \omega_d} \int_{H_1(0)} J(|\xi|) \frac{F'_1(\sqrt{s_\xi} S_\xi(u))}{\sqrt{s_\xi}} e_\xi \otimes \nabla^2 \bar{\omega}_\xi(x) d\xi \\ &\quad + \frac{2}{\epsilon \omega_d} \int_{H_1(0)} J(|\xi|) F''_1(\sqrt{s_\xi} S_\xi(u)) e_\xi \otimes \nabla \bar{\omega}_\xi(x) \otimes S_\xi(\nabla u) d\xi \\ &= h_1(u)(x) + h_2(u)(x) + h_3(u)(x) + h_4(u)(x) + h_5(u)(x) \end{aligned} \tag{63}$$

where we denote first, second, third, fourth, and fifth terms as $h_1, h_2, h_3, h_4,$ and $h_5,$ respectively. Estimating $\|P(u) - P(v)\|$ and $\|P(u) - P(v)\|_\infty$. From (61), we have

$$\begin{aligned}
 & |P(u)(x) - P(v)(x)| \\
 & \leq \frac{2}{\epsilon\omega_d} \int_{H_1(0)} J(|\xi|) \frac{1}{\sqrt{s_\xi}} |F'_1(\sqrt{s_\xi}S_\xi(u)) - F'_1(\sqrt{s_\xi}S_\xi(v))| d\xi \\
 & \leq \frac{2}{\epsilon\omega_d} \left(\sup_r |F'_1(r)| \right) \int_{H_1(0)} J(|\xi|) \frac{1}{\sqrt{s_\xi}} |\sqrt{s_\xi}S_\xi(u) - \sqrt{s_\xi}S_\xi(v)| d\xi \\
 & = \frac{2C_2}{\epsilon\omega_d} \int_{H_1(0)} J(|\xi|) |S_\xi(u) - S_\xi(v)| d\xi, \tag{64}
 \end{aligned}$$

where we used the fact that $|\bar{\omega}_\xi(x)| \leq 1$ and $|F'_1(r_1) - F'_1(r_2)| \leq C_2|r_1 - r_2|$. Since

$$|S_\xi(u) - S_\xi(v)| \leq \frac{|u(x + \epsilon\xi) - v(x + \epsilon\xi)| + |u(x) - v(x)|}{\epsilon|\xi|}$$

we have

$$\|P(u) - P(v)\|_\infty \leq \frac{2C_2}{\epsilon\omega_d} \int_{H_1(0)} J(|\xi|) \frac{2\|u - v\|_\infty}{\epsilon|\xi|} d\xi = \frac{L_1}{\epsilon^2} \|u - v\|_W \tag{65}$$

where we let $L_1 := 4C_2\bar{J}_1$.

From (64), we have

$$\begin{aligned}
 & \|P(u) - P(v)\|^2 \\
 & \leq \int_D \left(\frac{2C_2}{\epsilon\omega_d} \right)^2 \int_{H_1(0)} \int_{H_1(0)} \frac{J(|\xi|)}{|\xi|} \frac{J(|\eta|)}{|\eta|} |\xi| |S_\xi(u) - S_\xi(v)| |\eta| |S_\eta(u) - S_\eta(v)| \\
 & \quad d\xi d\eta dx.
 \end{aligned}$$

Using the identities $|a||b| \leq |a|^2/2 + |b|^2/2$ and $(a + b)^2 \leq 2a^2 + 2b^2$, we get

$$\begin{aligned}
 & \|P(u) - P(v)\|^2 \\
 & \leq \int_D \left(\frac{2C_2}{\epsilon\omega_d} \right)^2 \int_{H_1(0)} \int_{H_1(0)} \frac{J(|\xi|)}{|\xi|} \frac{J(|\eta|)}{|\eta|} \frac{|\xi|^2 |S_\xi(u) - S_\xi(v)|^2 + |\eta|^2 |S_\eta(u) - S_\eta(v)|^2}{2} d\xi d\eta dx \\
 & = 2 \int_D \left(\frac{2C_2}{\epsilon\omega_d} \right)^2 \int_{H_1(0)} \int_{H_1(0)} \frac{J(|\xi|)}{|\xi|} \frac{J(|\eta|)}{|\eta|} \frac{|\xi|^2 |S_\xi(u) - S_\xi(v)|^2}{2} d\xi d\eta dx \\
 & = \int_D \left(\frac{2C_2}{\epsilon\omega_d} \right)^2 \omega_d \bar{J}_1 \int_{H_1(0)} \frac{J(|\xi|)}{|\xi|} |\xi|^2 \frac{2|u(x + \epsilon\xi) - v(x + \epsilon\xi)|^2 + 2|u(x) - v(x)|^2}{\epsilon^2 |\xi|^2} d\xi dx \\
 & = \left(\frac{2C_2}{\epsilon\omega_d} \right)^2 \omega_d \bar{J}_1 \int_{H_1(0)} \frac{J(|\xi|)}{|\xi|} \frac{1}{\epsilon^2} \left[2 \int_D (|u(x + \epsilon\xi) - v(x + \epsilon\xi)|^2 + |u(x) - v(x)|^2) dx \right] d\xi \\
 & \leq \left(\frac{2C_2}{\epsilon\omega_d} \right)^2 \omega_d \bar{J}_1 \int_{H_1(0)} \frac{J(|\xi|)}{|\xi|} \frac{1}{\epsilon^2} [4\|u - v\|^2] d\xi, \tag{66}
 \end{aligned}$$

where we used symmetry wrt ξ and η in second equation. This gives

$$\|P(u) - P(v)\| \leq \frac{L_1}{\epsilon^2} \|u - v\| \leq \frac{L_1}{\epsilon^2} \|u - v\|_W. \quad (67)$$

Estimating $\|\nabla P(u) - \nabla P(v)\|$. From (62), we have

$$\|\nabla P(u) - \nabla P(v)\| \leq \|g_1(u) - g_1(v)\| + \|g_2(u) - g_2(v)\|.$$

Using $|\bar{\omega}_\xi(x)| \leq 1$, we get

$$\begin{aligned} & |g_1(u)(x) - g_1(v)(x)| \\ & \leq \frac{2}{\epsilon\omega_d} \int_{H_1(0)} J(|\xi|) |F_1''(\sqrt{s_\xi} S_\xi(u)) \nabla S_\xi(u) - F_1''(\sqrt{s_\xi} S_\xi(v)) \nabla S_\xi(v)| d\xi \\ & \leq \frac{2}{\epsilon\omega_d} \int_{H_1(0)} J(|\xi|) |F_1''(\sqrt{s_\xi} S_\xi(u)) - F_1''(\sqrt{s_\xi} S_\xi(v))| |\nabla S_\xi(u)| d\xi \\ & \quad + \frac{2}{\epsilon\omega_d} \int_{H_1(0)} J(|\xi|) |F_1''(\sqrt{s_\xi} S_\xi(v))| |\nabla S_\xi(u) - \nabla S_\xi(v)| d\xi \\ & \leq \frac{2C_3}{\epsilon\omega_d} \int_{H_1(0)} J(|\xi|) \sqrt{s_\xi} |S_\xi(u) - S_\xi(v)| |\nabla S_\xi(u)| d\xi \\ & \quad + \frac{2C_2}{\epsilon\omega_d} \int_{H_1(0)} J(|\xi|) |\nabla S_\xi(u) - \nabla S_\xi(v)| d\xi \\ & = I_1(x) + I_2(x) \end{aligned} \quad (68)$$

where we denote first and second term as $I_1(x)$ and $I_2(x)$. Proceeding similar to ((66)), we can show

$$\begin{aligned} \|I_1\|^2 &= \int_D \left(\frac{2C_3}{\epsilon\omega_d} \right)^2 \int_{H_1(0)} \int_{H_1(0)} \frac{J(|\xi|)}{|\xi|^{3/2}} \frac{J(|\eta|)}{|\eta|^{3/2}} |\xi|^{3/2} |\eta|^{3/2} \sqrt{s_\xi} \sqrt{s_\eta} \\ & \quad \times |S_\xi(u) - S_\xi(v)| |\nabla S_\xi(u)| |S_\eta(u) - S_\eta(v)| |\nabla S_\eta(u)| d\xi d\eta dx \\ & \leq \int_D \left(\frac{2C_3}{\epsilon\omega_d} \right)^2 \omega_d \bar{J}_{3/2} \int_{H_1(0)} \frac{J(|\xi|)}{|\xi|^{3/2}} |\xi|^{3/2} s_\xi |S_\xi(u) - S_\xi(v)|^2 |\nabla S_\xi(u)|^2 d\xi dx. \end{aligned} \quad (69)$$

Now

$$\int_D |S_\xi(u) - S_\xi(v)|^2 |\nabla S_\xi(u)|^2 dx$$

$$\begin{aligned} &\leq \frac{4\|u - v\|_\infty^2}{\epsilon^2|\xi|^2} \frac{1}{\epsilon^2|\xi|^2} \int_D 2(|\nabla u(x + \epsilon\xi)|^2 + |\nabla u(x)|^2) dx \\ &\leq \frac{16\|\nabla u\|^2\|u - v\|_\infty^2}{\epsilon^4|\xi|^4} \leq \frac{16\|u\|_W^2}{\epsilon^4|\xi|^4} \|u - v\|_W^2. \end{aligned}$$

Substituting above in (69) to get

$$\begin{aligned} \|I_1\|^2 &\leq \left(\frac{2C_3}{\epsilon\omega_d}\right)^2 \omega_d \bar{J}_{3/2} \int_{H_1(0)} \frac{J(|\xi|)}{|\xi|^{3/2}} |\xi|^3 \epsilon |\xi| \frac{16\|u\|_W^2}{\epsilon^4|\xi|^4} \|u - v\|_W^2 d\xi \\ &= \left(\frac{8C_3\bar{J}_{3/2}\|u\|_W}{\epsilon^{5/2}}\right)^2 \|u - v\|_W^2. \end{aligned}$$

Let $L_2 = 8C_3\bar{J}_{3/2}$ to write

$$\|I_1\| \leq \frac{L_2(\|u\|_W + \|v\|_W)}{\epsilon^{5/2}} \|u - v\|_W. \tag{70}$$

Similarly

$$\begin{aligned} \|I_2\|^2 &= \int_D \left(\frac{2C_2}{\epsilon\omega_d}\right)^2 \int_{H_1(0)} \int_{H_1(0)} \frac{J(|\xi|)}{|\xi|} \frac{J(|\eta|)}{|\eta|} |\xi||\eta| \\ &\quad \times |\nabla S_\xi(u) - \nabla S_\xi(v)| |\nabla S_\eta(u) - \nabla S_\eta(v)| d\xi d\eta dx \\ &\leq \left(\frac{2C_2}{\epsilon\omega_d}\right)^2 \omega_d \bar{J}_1 \int_{H_1(0)} \frac{J(|\xi|)}{|\xi|} |\xi|^2 \left[\int_D |\nabla S_\xi(u) - \nabla S_\xi(v)|^2 dx \right] d\xi. \end{aligned}$$

This gives

$$\|I_2\| \leq \frac{4C_2\bar{J}_1}{\epsilon^2} \|u - v\|_W = \frac{L_1}{\epsilon^2} \|u - v\|_W. \tag{71}$$

Thus

$$\|g_1(u) - g_1(v)\| \leq \frac{\sqrt{\epsilon}L_1 + L_2(\|u\|_W + \|v\|_W)}{\epsilon^{5/2}} \|u - v\|_W. \tag{72}$$

We now work on $|g_2(u)(x) - g_2(v)(x)|$ (see (62)). Noting the bound on $\nabla\bar{\omega}_\xi$, we get

$$|g_2(u)(x) - g_2(v)(x)|$$

$$\begin{aligned}
 &= \left| \frac{2}{\epsilon \omega_d} \int_{H_1(0)} J(|\xi|) \left[\frac{F'_1(\sqrt{s_\xi} S_\xi(u))}{\sqrt{s_\xi}} - \frac{F'_1(\sqrt{s_\xi} S_\xi(v))}{\sqrt{s_\xi}} \right] e_\xi \otimes \nabla \bar{\omega}_\xi(s) d\xi \right| \\
 &\leq \frac{2C_{\omega_1}}{\epsilon \omega_d} \int_{H_1(0)} J(|\xi|) \left| \frac{F'_1(\sqrt{s_\xi} S_\xi(u))}{\sqrt{s_\xi}} - \frac{F'_1(\sqrt{s_\xi} S_\xi(v))}{\sqrt{s_\xi}} \right| d\xi \\
 &\leq \frac{2C_{\omega_1} C_2}{\epsilon \omega_d} \int_{H_1(0)} J(|\xi|) |S_\xi(u) - S_\xi(v)| d\xi. \tag{73}
 \end{aligned}$$

Above is similar to (64) and therefore we get

$$\|g_2(u) - g_2(v)\| \leq \frac{4C_{\omega_1} C_2 \bar{J}_1}{\epsilon^2} \|u - v\|_W = \frac{C_{\omega_1} L_1}{\epsilon^2} \|u - v\|_W. \tag{74}$$

Combining (72) and (74) to write

$$\|\nabla P(u) - \nabla P(v)\| \leq \frac{\sqrt{\epsilon}(1 + C_{\omega_1})L_1 + L_2(\|u\|_W + \|v\|_W)}{\epsilon^{5/2}} \|u - v\|_W. \tag{75}$$

Estimating $\|\nabla^2 P(u) - \nabla^2 P(v)\|$. From (63), we have

$$\begin{aligned}
 &\|\nabla^2 P(u) - \nabla^2 P(v)\| \\
 &\leq \|h_1(u) - h_1(v)\| + \|h_2(u) - h_2(v)\| + \|h_3(u) - h_3(v)\| \\
 &\quad + \|h_4(u) - h_4(v)\| + \|h_5(u) - h_5(v)\|. \tag{76}
 \end{aligned}$$

We can show, using the fact $|\bar{\omega}_\xi(x)| \leq 1$ and $|F''_1(r_1) - F''_1(r_2)| \leq C_3|r_1 - r_2|$, that

$$\begin{aligned}
 |h_1(u)(x) - h_1(v)(x)| &\leq \frac{2C_3}{\epsilon \omega_d} \int_{H_1(0)} J(|\xi|) \sqrt{s_\xi} |S_\xi(u) - S_\xi(v)| |S_\xi(\nabla^2 u)| d\xi \\
 &\quad + \frac{2C_2}{\epsilon \omega_d} \int_{H_1(0)} J(|\xi|) |S_\xi(\nabla^2 u) - S_\xi(\nabla^2 v)| d\xi \\
 &= I_3(x) + I_4(x). \tag{77}
 \end{aligned}$$

Following similar steps used above, we can show

$$\|I_3\| \leq \frac{8C_3 \bar{J}_{3/2} \|u\|_W}{\epsilon^{5/2}} \|u - v\|_W \leq \frac{L_2(\|u\|_W + \|v\|_W)}{\epsilon^{5/2}} \|u - v\|_W \tag{78}$$

and

$$\|I_4\| \leq \frac{4C_2 \bar{J}_1}{\epsilon^2} \|u - v\|_W = \frac{L_1}{\epsilon^2} \|u - v\|_W, \tag{79}$$

where $L_1 = 4C_2\bar{J}_1, L_2 = 8C_3\bar{J}_{3/2}$.

Next we focus on $|h_2(u)(x) - h_2(v)(x)|$ and get

$$\begin{aligned}
 & |h_2(u)(x) - h_2(v)(x)| \\
 & \leq \frac{2}{\epsilon\omega_d} \int_{H_1(0)} J(|\xi|)\sqrt{s_\xi}|F_1'''(\sqrt{s_\xi}S_\xi(u)) - F_1'''(\sqrt{s_\xi}S_\xi(v))||S_\xi(\nabla u)|^2 d\xi \\
 & + \frac{2}{\epsilon\omega_d} \int_{H_1(0)} J(|\xi|)\sqrt{s_\xi}|F_1'''(\sqrt{s_\xi}S_\xi(v))||S_\xi(\nabla u) \otimes S_\xi(\nabla u) \\
 & - S_\xi(\nabla v) \otimes S_\xi(\nabla v)|d\xi \leq \frac{2C_4}{\epsilon\omega_d} \int_{H_1(0)} J(|\xi|)s_\xi|S_\xi(u) - S_\xi(v)||S_\xi(\nabla u)|^2 d\xi \\
 & + \frac{2C_3}{\epsilon\omega_d} \int_{H_1(0)} J(|\xi|)\sqrt{s_\xi}|S_\xi(\nabla u) \otimes S_\xi(\nabla u) - S_\xi(\nabla v) \otimes S_\xi(\nabla v)|d\xi \\
 & = I_5(x) + I_6(x). \tag{80}
 \end{aligned}$$

Proceeding as below for $||I_5||^2$

$$\begin{aligned}
 & ||I_5||^2 \\
 & \leq \int_D \left(\frac{2C_4}{\epsilon\omega_d}\right)^2 \int_{H_1(0)} \int_{H_1(0)} \frac{J(|\xi|)}{|\xi|^2} \frac{J(|\eta|)}{|\eta|^2} |\xi|^2 s_\xi |\eta|^2 s_\eta \\
 & \times |S_\xi(u) - S_\xi(v)||S_\xi(\nabla u)|^2 |S_\eta(u) - S_\eta(v)||S_\eta(\nabla u)|^2 d\xi d\eta dx \\
 & \leq \int_D \left(\frac{2C_4}{\epsilon\omega_d}\right)^2 \omega_d \bar{J}_2 \int_{H_1(0)} \frac{J(|\xi|)}{|\xi|^2} |\xi|^4 s_\xi^2 |S_\xi(u) - S_\xi(v)|^2 |S_\xi(\nabla u)|^4 d\xi dx \\
 & \leq \left(\frac{2C_4}{\epsilon\omega_d}\right)^2 \omega_d \bar{J}_2 \int_{H_1(0)} \frac{J(|\xi|)}{|\xi|^2} |\xi|^4 s_\xi^2 \frac{4||u - v||_\infty^2}{\epsilon^2 |\xi|^2} \left[\int_D |S_\xi(\nabla u)|^4 dx \right] d\xi. \tag{81}
 \end{aligned}$$

We estimate the term in square bracket. Using the identity $(|a| + |b|)^4 \leq (2|a|^2 + 2|b|^2)^2 \leq 8|a|^4 + 8|b|^4$, we have

$$\begin{aligned}
 \int_D |S_\xi(\nabla u)|^4 dx & \leq \frac{8}{\epsilon^4 |\xi|^4} \int_D (|\nabla u(x + \epsilon\xi)|^4 + |\nabla u(x)|^4) dx \\
 & \leq \frac{16}{\epsilon^4 |\xi|^4} ||\nabla u||_{L^4(D; \mathbb{R}^{d \times d})}^4. \tag{82}
 \end{aligned}$$

where $||u||_{L^4(D; \mathbb{R}^d)} = [\int_D |u|^4 dx]^{1/4}$. Using Sobolev embedding property of $u \in H_0^2(D; \mathbb{R}^d)$ as mentioned in (51), we get

$$\int_D |S_\xi(\nabla u)|^4 dx \leq \frac{16}{\epsilon^4 |\xi|^4} C_e^4 \|\nabla u\|_{H^1(D; \mathbb{R}^{d \times d})}^4 \leq \frac{16 C_e^4}{\epsilon^4 |\xi|^4} \|u\|_W^4. \quad (83)$$

Substituting above to get

$$\|I_5\|^2 \leq \left(\frac{2C_4}{\epsilon \omega_d} \right)^2 \omega_d \bar{J}_2 \int_{H_1(0)} \frac{J(|\xi|)}{|\xi|^2} |\xi|^{4S_\xi^2} \frac{4\|u-v\|_\infty^2}{\epsilon^2 |\xi|^2} \frac{16C_e^4}{\epsilon^4 |\xi|^4} \|u\|_W^4 d\xi$$

Above gives

$$\|I_5\| \leq \frac{16C_4 C_e^2 \bar{J}_2 \|u\|_W^2}{\epsilon^3} \|u-v\|_W \leq \frac{L_3 (\|u\|_W + \|v\|_W)^2}{\epsilon^3} \|u-v\|_W \quad (84)$$

where we let $L_3 = 16C_4 C_e^2 \bar{J}_2$.

Next, using

$$|S_\xi(\nabla u) \otimes S_\xi(\nabla u) - S_\xi(\nabla v) \otimes S_\xi(\nabla v)| \leq (|S_\xi(\nabla u)| + |S_\xi(\nabla v)|) |S_\xi(\nabla u) - S_\xi(\nabla v)|$$

to estimate $\|I_6\|$ as follows:

$$\begin{aligned} & \|I_6\|^2 \\ & \leq \int_D \left(\frac{2C_3}{\epsilon \omega_d} \right)^2 \int_{H_1(0)} \int_{H_1(0)} \frac{J(|\xi|)}{|\xi|^{3/2}} \frac{J(|\eta|)}{|\eta|^{3/2}} |\xi|^{3/2} |\eta|^{3/2} \sqrt{S_\xi S_\eta} \\ & \quad \times (|S_\xi(\nabla u)| + |S_\xi(\nabla v)|) |S_\xi(\nabla u) - S_\xi(\nabla v)| \times (|S_\eta(\nabla u)| + |S_\eta(\nabla v)|) |S_\eta(\nabla u) \\ & \quad - S_\eta(\nabla v)| d\xi d\eta dx \leq \int_D \left(\frac{2C_3}{\epsilon \omega_d} \right)^2 \omega_d \bar{J}_{3/2} \int_{H_1(0)} \frac{J(|\xi|)}{|\xi|^{3/2}} |\xi|^3 \epsilon |\xi| (|S_\xi(\nabla u)| \\ & \quad + |S_\xi(\nabla v)|)^2 |S_\xi(\nabla u) - S_\xi(\nabla v)|^2 d\xi dx = \left(\frac{2C_3}{\epsilon \omega_d} \right)^2 \omega_d \bar{J}_{3/2} \int_{H_1(0)} \frac{J(|\xi|)}{|\xi|^{3/2}} |\xi|^3 \epsilon |\xi| \\ & \quad \left[\int_D (|S_\xi(\nabla u)| + |S_\xi(\nabla v)|)^2 |S_\xi(\nabla u) - S_\xi(\nabla v)|^2 dx \right] d\xi. \end{aligned} \quad (85)$$

We focus on the term in square bracket. Using Holder inequality, we have

$$\begin{aligned} & \int_D (|S_\xi(\nabla u)| + |S_\xi(\nabla v)|)^2 |S_\xi(\nabla u) - S_\xi(\nabla v)|^2 dx \\ & \leq \left(\int_D (|S_\xi(\nabla u)| + |S_\xi(\nabla v)|)^4 dx \right)^{1/2} \left(\int_D |S_\xi(\nabla u) - S_\xi(\nabla v)|^4 dx \right)^{1/2}. \end{aligned} \quad (86)$$

Using $(|a| + |b|)^4 \leq 8|a|^4 + 8|b|^4$, we get

$$\begin{aligned}
 \int_D (|S_\xi(\nabla u)| + |S_\xi(\nabla v)|)^4 dx &\leq 8 \left[\int_D |S_\xi(\nabla u)|^4 dx + \int_D |S_\xi(\nabla v)|^4 dx \right] \\
 &\leq 8 \left[\frac{8}{\epsilon^4 |\xi|^4} \int_D (|\nabla u(x + \epsilon \xi)|^4 + |\nabla u(x)|^4) dx + \frac{8}{\epsilon^4 |\xi|^4} \int_D (|\nabla v(x + \epsilon \xi)|^4 \right. \\
 &\quad \left. + |\nabla v(x)|^4) dx \right] \leq \frac{128}{\epsilon^4 |\xi|^4} (\|\nabla u\|_{L^4(D; \mathbb{R}^{d \times d})}^4 + \|\nabla v\|_{L^4(D; \mathbb{R}^{d \times d})}^4) \\
 &\leq \frac{128C_e^4}{\epsilon^4 |\xi|^4} (\|\nabla u\|_{H^1(D; \mathbb{R}^{d \times d})}^4 + \|\nabla v\|_{H^1(D; \mathbb{R}^{d \times d})}^4) \leq \frac{128C_e^4}{\epsilon^4 |\xi|^4} (\|u\|_W^4 + \|v\|_W^4) \\
 &\leq \frac{128C_e^4}{\epsilon^4 |\xi|^4} (\|u\|_W + \|v\|_W)^4. \tag{87}
 \end{aligned}$$

where we used Sobolev embedding property (51) in third last step. Proceeding similarly to get

$$\begin{aligned}
 &\int_D |S_\xi(\nabla u) - S_\xi(\nabla v)|^4 dx \\
 &\leq \frac{8}{\epsilon^4 |\xi|^4} \left[\int_D |\nabla(u - v)(x + \epsilon \xi)|^4 dx + \int_D |\nabla(u - v)(x)|^4 dx \right] \\
 &\leq \frac{16}{\epsilon^4 |\xi|^4} \|\nabla(u - v)\|_{L^4(D; \mathbb{R}^{d \times d})}^4 \\
 &\leq \frac{16C_e^4}{\epsilon^4 |\xi|^4} \|u - v\|_W^4. \tag{88}
 \end{aligned}$$

Substituting (87) and (88) into (86) to get

$$\begin{aligned}
 &\int_D (|S_\xi(\nabla u)| + |S_\xi(\nabla v)|)^2 |S_\xi(\nabla u) - S_\xi(\nabla v)|^2 dx \\
 &\leq \left(\frac{128C_e^4}{\epsilon^4 |\xi|^4} (\|u\|_W + \|v\|_W)^4 \right)^{1/2} \left(\frac{16C_e^4}{\epsilon^4 |\xi|^4} \|u - v\|_W^4 \right)^{1/2} \\
 &= \frac{32\sqrt{2}C_e^4}{\epsilon^4 |\xi|^4} (\|u\|_W + \|v\|_W)^2 \|u - v\|_W^2 \\
 &\leq \frac{64C_e^4}{\epsilon^4 |\xi|^4} (\|u\|_W + \|v\|_W)^2 \|u - v\|_W^2.
 \end{aligned}$$

Substituting above in (85) to get

$$\begin{aligned} & \|I_6\|^2 \\ & \leq \left(\frac{2C_3}{\epsilon\omega_d}\right)^2 \omega_d \bar{J}_{3/2} \int_{H_1(0)} \frac{J(|\xi|)}{|\xi|^{3/2}} |\xi|^3 \epsilon |\xi| \left[\frac{64C_e^4}{\epsilon^4 |\xi|^4} (\|u\|_W + \|v\|_W)^2 \|u - v\|_W^2 \right] d\xi. \end{aligned}$$

From above we have

$$\|I_6\| \leq \frac{16C_3C_e^2 \bar{J}_{3/2} (\|u\|_W + \|v\|_W)}{\epsilon^{5/2}} \|u - v\|_W = \frac{L_4 (\|u\|_W + \|v\|_W)}{\epsilon^{5/2}} \|u - v\|_W, \tag{89}$$

where we let $L_4 = 16C_3C_e^2 \bar{J}_{3/2}$.

From the expression of $h_3(u)(x)$ and $h_5(u)(x)$, we find that it is similar to term $g_1(u)(x)$ from the point of view of L^2 norm. Also, $h_4(u)(x)$ is similar to $P(u)(x)$. We easily have

$$|h_4(u)(x) - h_4(v)(x)| \leq \frac{2C_2C_{\omega_2}}{\epsilon\omega_d} \int_{H_1(0)} J(|\xi|) |S_\xi(u) - S_\xi(v)| d\xi,$$

where we used the fact that $|\nabla^2 \bar{\omega}_\xi(x)| \leq C_{\omega_2}$. Above is similar to the bound on $|P(u)(x) - P(v)(x)|$ (see (64)); therefore we have

$$\|h_4(u) - h_4(v)\| \leq \frac{L_1 C_{\omega_2}}{\epsilon^2} \|u - v\|_W. \tag{90}$$

Similarly, we have

$$\begin{aligned} |h_3(u)(x) - h_3(v)(x)| & \leq \frac{2}{\epsilon\omega_d} \int_{H_1(0)} J(|\xi|) |F_1''(\sqrt{s_\xi} S_\xi(u)) \\ & - F_1''(\sqrt{s_\xi} S_\xi(v))| |\nabla \bar{\omega}_\xi(x)| |S_\xi(\nabla u)| d\xi + \frac{2}{\epsilon\omega_d} \int_{H_1(0)} J(|\xi|) |F_1''(\sqrt{s_\xi} S_\xi(v))| |e_\xi \\ & \otimes \nabla \bar{\omega}_\xi(x) \otimes S_\xi(\nabla u) - e_\xi \otimes \nabla \bar{\omega}_\xi(x) \otimes S_\xi(\nabla v)| d\xi \leq \frac{2C_3C_{\omega_1}}{\epsilon\omega_d} \int_{H_1(0)} \\ & J(|\xi|) \sqrt{s_\xi} |S_\xi(u) - S_\xi(v)| |S_\xi(\nabla u)| d\xi + \frac{2C_2C_{\omega_1}}{\epsilon\omega_d} \int_{H_1(0)} J(|\xi|) |S_\xi(\nabla u) \\ & - S_\xi(\nabla v)| d\xi = C_{\omega_1} (I_1(x) + I_2(x)), \end{aligned} \tag{91}$$

where $I_1(x)$ and $I_2(x)$ are given by (68). From (70) to (71), we have

$$\begin{aligned} \|h_3(u) - h_3(v)\| &\leq C_{\omega_1}(\|I_1\| + \|I_2\|) \\ &\leq \frac{\sqrt{\epsilon}C_{\omega_1}L_1 + C_{\omega_1}L_2(\|u\|_W + \|v\|_W)}{\epsilon^{5/2}}\|u - v\|_W. \end{aligned} \tag{92}$$

Expression of $h_3(u)$ and $h_5(u)$ is similar and hence we have

$$\begin{aligned} \|h_5(u) - h_5(v)\| &\leq C_{\omega_1}(\|I_1\| + \|I_2\|) \\ &\leq \frac{\sqrt{\epsilon}C_{\omega_1}L_1 + C_{\omega_1}L_2(\|u\|_W + \|v\|_W)}{\epsilon^{5/2}}\|u - v\|_W. \end{aligned} \tag{93}$$

Collecting our results delivers the bound

$$\begin{aligned} &\|\nabla^2 P(u) - \nabla^2 P(v)\| \\ &\leq \left[\frac{\epsilon L_1 + \sqrt{\epsilon}L_2(\|u\|_W + \|v\|_W) + L_3(\|u\|_W + \|v\|_W)^2 + \sqrt{\epsilon}L_4(\|u\|_W + \|v\|_W)}{\epsilon^3} \right. \\ &\quad \left. + \frac{\epsilon C_{\omega_2}L_1 + 2\epsilon C_{\omega_1}L_1 + 2\sqrt{\epsilon}C_{\omega_1}L_2(\|u\|_W + \|v\|_W)}{\epsilon^3} \right] \|u - v\|_W \\ &\leq \left[\frac{\epsilon(1 + 2C_{\omega_1} + C_{\omega_2})L_1 + \sqrt{\epsilon}(L_2 + 2C_{\omega_1}L_2 + L_4)(\|u\|_W + \|v\|_W)}{\epsilon^3} \right. \\ &\quad \left. + \frac{L_3(\|u\|_W + \|v\|_W)^2}{\epsilon^3} \right] \|u - v\|_W. \end{aligned} \tag{94}$$

We now combine (65), (67), (75), and (94) to get

$$\begin{aligned} &\|P(u) - P(v)\|_W \\ &\leq \left[\frac{2\epsilon L_1 + \epsilon(1 + C_{\omega_1})L_1 + \sqrt{\epsilon}(\|u\|_W + \|v\|_W)}{\epsilon^3} \right. \\ &\quad \left. + \frac{\epsilon(1 + 2C_{\omega_1} + C_{\omega_2})L_1 + \sqrt{\epsilon}(L_2 + 2C_{\omega_1}L_2 + L_4)(\|u\|_W + \|v\|_W)}{\epsilon^3} \right. \\ &\quad \left. + \frac{L_3(\|u\|_W + \|v\|_W)^2}{\epsilon^3} \right] \|u - v\|_W. \end{aligned} \tag{95}$$

Finally we let

$$\bar{L}_1 := (4 + 3C_{\omega_1} + C_{\omega_2})L_1, \quad \bar{L}_2 := (1 + 2C_{\omega_1})L_2 + L_4, \quad \bar{L}_3 := L_3 \tag{96}$$

and write

$$\begin{aligned} & \|P(u) - P(v)\|_W \\ & \leq \frac{\bar{L}_1 + \bar{L}_2(\|u\|_W + \|v\|_W) + \bar{L}_3(\|u\|_W + \|v\|_W)^2}{\epsilon^3} \|u - v\|_W. \end{aligned} \quad (97)$$

This completes the proof of (55).

We now obtain an upper bound on the peridynamic force. Note that $F'_1(0) = 0$ and $S_\xi(v) = 0$ if $v = 0$. Substituting $v = 0$ in (65) and (67) to get

$$\|P(u)\| + \|P(u)\|_\infty \leq \frac{2L_1}{\epsilon^2} \|u\|_W. \quad (98)$$

For $\|g_1(u)\|$ and $\|g_2(u)\|$ we proceed differently. For $\|g_2(u)\|$, we substitute $v = 0$ in (74) to get

$$\|g_2(u)\| \leq \frac{C_{\omega_1} L_1}{\epsilon^2} \|u\|_W. \quad (99)$$

To estimate $\|g_1(u)\|$, we first estimate

$$\begin{aligned} |g_1(u)(x)| & \leq \frac{2C_2}{\epsilon\omega_d} \int_{H_1(0)} J(|\xi|) |\nabla S_\xi(u)| d\xi \\ & \leq \frac{2C_2}{\epsilon^2\omega_d} \int_{H_1(0)} \frac{J(|\xi|)}{|\xi|} (|\nabla u(x + \epsilon\xi)| + |\nabla u(x)|) d\xi, \end{aligned} \quad (100)$$

and we obtain

$$\begin{aligned} \|g_1(u)\|^2 & \leq \left(\frac{2C_2}{\epsilon^2\omega_d}\right)^2 \omega_d \bar{J}_1 \int_{H_1(0)} \frac{J(|\xi|)}{|\xi|} \left[\int_D (|\nabla u(x + \epsilon\xi)| + |\nabla u(x)|)^2 dx \right] d\xi \\ & \leq \left(\frac{4C_2\bar{J}_1}{\epsilon^2}\right)^2 \|\nabla u\|^2 \end{aligned} \quad (101)$$

i.e.

$$\|g_1(u)\| \leq \frac{L_1}{\epsilon^2} \|u\|_W. \quad (102)$$

Combining (99) and (102) gives

$$\|\nabla P(u)\| \leq \frac{(1 + C_{\omega_1})L_1}{\epsilon^2} \|u\|_W. \quad (103)$$

We need to estimate $\|\nabla^2 P(u)\|$. We have from (63)

$$\|\nabla^2 P(u)\| \leq \|h_1(u)\| + \|h_2(u)\| + \|h_3(u)\| + \|h_4(u)\| + \|h_5(u)\|.$$

From the expression of $h_1(u)$ and $h_2(u)$, we find that

$$\|h_1(u)\| \leq \frac{4C_2\bar{J}_1}{\epsilon^2} \|u\|_W = \frac{L_1}{\epsilon^2} \|u\|_W \quad \text{and} \quad \|h_2(u)\| \leq \frac{8C_3C_e^2\bar{J}_{3/2}}{\epsilon^{5/2}} \|u\|_W^2 \leq \frac{L_4}{\epsilon^{5/2}} \|u\|_W^2,$$

where $L_4 = 16C_3C_e^2\bar{J}_{3/2}$. Case of $\|h_3(u)\|$ and $\|h_5(u)\|$ is similar to $\|g_1(u)\|$, and case of $\|h_4(u)\|$ is similar to $\|P(u)\|$.

$$\|h_4(u)\| \leq \frac{C_{\omega_2}L_1}{\epsilon^2} \|u\|_W$$

and

$$\|h_3(u)\| \leq \frac{C_{\omega_1}L_1}{\epsilon^2} \|u\|_W \quad \text{and} \quad \|h_5(u)\| \leq \frac{C_{\omega_1}L_1}{\epsilon^2} \|u\|_W.$$

We combine the inequalities listed above to get

$$\|\nabla^2 P(u)\| \leq \frac{\sqrt{\epsilon}(1 + C_{\omega_2} + 2C_{\omega_1})L_1 + L_4\|u\|_W}{\epsilon^{5/2}} \|u\|_W. \tag{104}$$

Finally, after combining (98), (103), and (104), we get

$$\|P(u)\|_W \leq \frac{\sqrt{\epsilon}(4 + 3C_{\omega_1} + C_{\omega_2})L_1 + L_4\|u\|_W}{\epsilon^{5/2}} \|u\|_W.$$

We let

$$\bar{L}_4 := \bar{L}_1 \quad \text{and} \quad \bar{L}_5 := L_4 \tag{105}$$

to write

$$\|P(u)\|_W \leq \frac{\bar{L}_4\|u\|_W + \bar{L}_5\|u\|_W^2}{\epsilon^{5/2}}. \tag{106}$$

This completes the proof of (56) and this completes the proof of Theorem 3.

Local and Global Existence of Solution in $H^2 \cap L^\infty$ Space

We now prove Theorem 4. We first prove local existence for a finite time interval. We find that we can choose this time interval independent of the initial data. We

repeat the local existence theorem to uniquely continue the local solution over any finite time interval. The existence and uniqueness of local solutions is stated in the following theorem.

Theorem 5 (Local existence and uniqueness). *Given $X = W \times W$, $b(t) \in W$, and initial data $x_0 = (u_0, v_0) \in X$. We suppose that $b(t)$ is continuous in time over some time interval $I_0 = (-T, T)$ and satisfies $\sup_{t \in I_0} \|b(t)\|_W < \infty$. Then, there exists a time interval $I' = (-T', T') \subset I_0$ and unique solution $y = (y^1, y^2)$ such that $y \in C^1(I'; X)$ and*

$$y(t) = x_0 + \int_0^t F^\epsilon(y(\tau), \tau) d\tau, \text{ for } t \in I' \tag{107}$$

or equivalently

$$y'(t) = F^\epsilon(y(t), t), \text{ with } y(0) = x_0, \text{ for } t \in I'$$

where $y(t)$ and $y'(t)$ are Lipschitz continuous in time for $t \in I' \subset I_0$.

Proof. To prove Theorem 5, we proceed as follows. Write $y(t) = (y^1(t), y^2(t))$ with $\|y\|_X = \|y^1(t)\|_W + \|y^2(t)\|_W$. Let us consider $R > \|x_0\|_X$ and define the ball $B(0, R) = \{y \in X : \|y\|_X < R\}$. Let $r < \min\{1, R - \|x_0\|_X\}$. We clearly have $r^2 < (R - \|x_0\|_X)^2$ as well as $r^2 < r < R - \|x_0\|_X$. Consider the ball $B(x_0, r^2)$ defined by

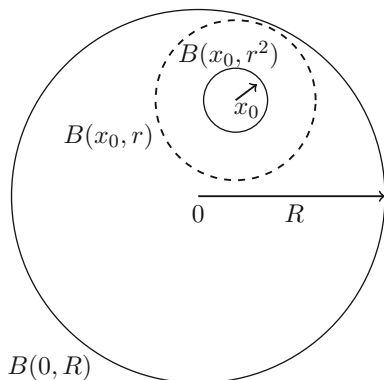
$$B(x_0, r^2) = \{y \in X : \|y - x_0\|_X < r^2\}. \tag{108}$$

Then we have $B(x_0, r^2) \subset B(x_0, r) \subset B(0, R)$ (see Fig. 7).

To recover the existence and uniqueness, we introduce the transformation

$$S_{x_0}(y)(t) = x_0 + \int_0^t F^\epsilon(y(\tau), \tau) d\tau.$$

Fig. 7 Geometry



Introduce $0 < T' < T$ and the associated set $Y(T')$ of functions in W taking values in $B(x_0, r^2)$ for $I' = (-T', T') \subset I_0 = (-T, T)$. The goal is to find appropriate interval $I' = (-T', T')$ for which S_{x_0} maps into the corresponding set $Y(T')$. Writing out the transformation with $y(t) \in Y(T')$ gives

$$S_{x_0}^1(y)(t) = x_0^1 + \int_0^t y^2(\tau) d\tau \tag{109}$$

$$S_{x_0}^2(y)(t) = x_0^2 + \int_0^t (-\nabla PD^\epsilon(y^1(\tau)) + b(\tau)) d\tau. \tag{110}$$

We have from (109)

$$\|S_{x_0}^1(y)(t) - x_0^1\|_W \leq \sup_{t \in (-T', T')} \|y^2(t)\|_W T'. \tag{111}$$

Using bound on $-\nabla PD^\epsilon$ in Theorem 3, we have from (110)

$$\|S_{x_0}^2(y)(t) - x_0^2\|_W \leq \int_0^t \left[\frac{\bar{L}_4}{\epsilon^{5/2}} \|y^1(\tau)\|_W + \frac{\bar{L}_5}{\epsilon^{5/2}} \|y^1(\tau)\|_W^2 + \|b(\tau)\|_W \right] d\tau. \tag{112}$$

Let $\bar{b} = \sup_{t \in I_0} \|b(t)\|_W$. Noting that transformation S_{x_0} is defined for $t \in I' = (-T', T')$ and $y(\tau) = (y^1(\tau), y^2(\tau)) \in B(x_0, r^2) \subset B(0, R)$ as $y \in Y(T')$, we have from (112) and (111)

$$\begin{aligned} \|S_{x_0}^1(y)(t) - x_0^1\|_W &\leq RT', \\ \|S_{x_0}^2(y)(t) - x_0^2\|_W &\leq \left[\frac{\bar{L}_4 R + \bar{L}_5 R^2}{\epsilon^{5/2}} + \bar{b} \right] T'. \end{aligned}$$

Adding gives

$$\|S_{x_0}(y)(t) - x_0\|_X \leq \left[\frac{\bar{L}_4 R + \bar{L}_5 R^2}{\epsilon^{5/2}} + R + \bar{b} \right] T'. \tag{113}$$

Choosing T' as below

$$T' < \frac{r^2}{\left[\frac{\bar{L}_4 R + \bar{L}_5 R^2}{\epsilon^{5/2}} + R + \bar{b} \right]} \tag{114}$$

will result in $S_{x_0}(y) \in Y(T')$ for all $y \in Y(T')$ as

$$\|S_{x_0}(y)(t) - x_0\|_X < r^2. \tag{115}$$

Since $r^2 < (R - \|x_0\|_X)^2$, we have

$$T' < \frac{r^2}{\left[\frac{\bar{L}_4 R + \bar{L}_5 R^2}{\epsilon^{5/2}} + R + \bar{b} \right]} < \frac{(R - \|x_0\|_X)^2}{\left[\frac{\bar{L}_4 R + \bar{L}_5 R^2}{\epsilon^{5/2}} + R + \bar{b} \right]}.$$

Let $\theta(R)$ be given by

$$\theta(R) := \frac{(R - \|x_0\|_X)^2}{\left[\frac{\bar{L}_4 R + \bar{L}_5 R^2}{\epsilon^{5/2}} + R + \bar{b} \right]}. \tag{116}$$

$\theta(R)$ is increasing with $R > 0$ and satisfies

$$\theta_\infty := \lim_{R \rightarrow \infty} \theta(R) = \frac{\epsilon^{5/2}}{\bar{L}_5}. \tag{117}$$

So given R and $\|x_0\|_X$, we choose T' according to

$$\frac{\theta(R)}{2} < T' < \theta(R), \tag{118}$$

and set $I' = (-T', T')$. This way we have shown that for time domain I' the transformation $S_{x_0}(y)(t)$ as defined in Eqs. 109 and 110 maps $Y(T')$ into itself. Existence and uniqueness of solution can be established using (Theorem 6.10, Driver 2003). \square

We now prove Theorem 1. From the proof of Theorem 2 above, we have a unique local solution over a time domain $(-T', T')$ with $\frac{\theta(R)}{2} < T'$. Since $\theta(R) \nearrow \epsilon^{5/2}/\bar{L}_5$ as $R \nearrow \infty$, we can fix a tolerance $\eta > 0$ so that $[(\epsilon^{5/2}/2\bar{L}_5) - \eta] > 0$. Then for any initial condition in W and $b = \sup_{t \in [-T, T]} \|b(t)\|_W$, we can choose R sufficiently large so that $\|x_0\|_X < R$ and $0 < (\epsilon^{5/2}/2\bar{L}_5) - \eta < T'$. Since choice of T' is independent of initial condition and R , we can always find local solutions for time intervals $(-T', T')$ for T' larger than $[(\epsilon^{5/2}/2\bar{L}_5) - \eta] > 0$. Therefore we apply the local existence and uniqueness result to uniquely continue local solutions up to an arbitrary time interval $(-T, T)$.

Conclusions: Convergence of Regular Solutions in the Limit of Vanishing Horizon

In this final section, we examine the behavior of bounded Hölder continuous solutions as the peridynamic horizon tends to zero. We find that the solutions converge to a limiting sharp fracture evolution with bounded Griffiths fracture energy and satisfy the linear elastic wave equation away from the fracture set. We look at a subset of Hölder solutions that are differentiable in the spatial

variables to show that sharp fracture evolutions can be approached by spatially smooth evolutions in the limit of vanishing nonlocality. As ϵ approaches zero, derivatives can become large but must localize to surfaces across which the limiting evolution jumps. These conclusions are reported in Jha and Lipton (2017b). The same behavior can be recovered for bounded H^2 solutions in the limit of vanishing horizon. These results support the numerical simulation for more regular nonlocal evolutions that approximate sharp fracture in the limit of vanishing nonlocality. In the next chapter we provide a priori estimates of the finite difference and finite element approximations to fracture evolution for nonlocal models with horizon $\epsilon > 0$.

To fix ideas consider a sequence of peridynamic horizons $\epsilon_k = 1/k, k = 1, \dots$ and the associated Hölder continuous solutions $\mathbf{u}^{\epsilon_k}(t, \mathbf{x})$ of the peridynamic initial value problem Eqs. 1, 2, and 3. We assume that the initial conditions $\mathbf{u}_0^{\epsilon_k}, \mathbf{v}_0^{\epsilon_k}$ have uniformly bounded peridynamic energy and mean square initial velocity given by

$$\sup_{\epsilon_k} PD^{\epsilon_k}(\mathbf{u}_0^{\epsilon_k}) < \infty \text{ and } \sup_{\epsilon_k} \|\mathbf{v}_0^{\epsilon_k}\|_{L^2(D; \mathbb{R}^d)} < \infty.$$

Moreover we suppose that $\mathbf{u}_0^{\epsilon_k}, \mathbf{v}_0^{\epsilon_k}$ are differentiable on D and that they converge in $L^2(D; \mathbb{R})$ to $\mathbf{u}_0^0, \mathbf{v}_0^0$ with bounded Griffith free energy given by

$$\int_D 2\mu |\mathcal{E}\mathbf{u}_0^0|^2 + \lambda |\text{div} \mathbf{u}_0^0|^2 dx + \mathcal{G}_c \mathcal{H}^{d-1}(J_{\mathbf{u}_0^0}) \leq C < \infty,$$

where $J_{\mathbf{u}_0^0}$ denotes an initial fracture surface given by the jumps in the initial deformation \mathbf{u}_0^0 and $\mathcal{H}^2(J_{\mathbf{u}_0^0(t)})$ is its two-dimensional Hausdorff measure of the jump set. Here $\mathcal{E}\mathbf{u}_0^0$ is the elastic strain and $\text{div} \mathbf{u}_0^0 = \text{Tr}(\mathcal{E}\mathbf{u}_0^0)$. The constants μ, λ are given by the explicit formulas

$$\text{and } \mu = \lambda = \frac{1}{5} f'(0) \int_0^1 r^d J(r) dr, \quad d = 2, 3$$

and

$$\mathcal{G}_c = \frac{3}{2} f_\infty \int_0^1 r^d J(r) dr, \quad d = 2, 3,$$

where $f'(0)$ and f_∞ are defined by Eq. 6. Here $\mu = \lambda$ and is a consequence of the central force model used in cohesive dynamics. Last we suppose as in Lipton (2016) that the solutions are uniformly bounded, i.e.,

$$\sup_{\epsilon_k} \sup_{[0, T]} \|\mathbf{u}^{\epsilon_k}(t)\|_{L^\infty(D; \mathbb{R}^d)} < \infty,$$

The Hölder solutions $\mathbf{u}^{\epsilon_k}(t, \mathbf{x})$ naturally belong to $L^2(D; \mathbb{R}^d)$ for all $t \in [0, T]$, and we can directly apply the Gronwall inequality (Equation (6.9) of Lipton

2016) together with Theorems 6.2 and 6.4 of Lipton (2016) to conclude similar to Theorems 5.1 and 5.2 of Lipton (2016) that there is at least one “cluster point” $\mathbf{u}^0(t, \mathbf{x})$ belonging to $C([0, T]; L^2(D; \mathbb{R}^d))$ and subsequence, also denoted by $\mathbf{u}^{\epsilon_k}(t, \mathbf{x})$ for which

$$\lim_{\epsilon_k \rightarrow 0} \max_{0 \leq t \leq T} \{ \|\mathbf{u}^{\epsilon_k}(t) - \mathbf{u}^0(t)\|_{L^2(D; \mathbb{R}^d)} \} = 0.$$

Moreover it follows from Lipton (2016) that the limit evolution $\mathbf{u}^0(t, \mathbf{x})$ has a weak derivative $\mathbf{u}_t^0(t, \mathbf{x})$ belonging to $L^2([0, T] \times D; \mathbb{R}^d)$. For each time $t \in [0, T]$, we can apply methods outlined in Lipton (2016) to find that the cluster point $\mathbf{u}^0(t, \mathbf{x})$ is a special function of bounded deformation (see Ambrosio et al. 1997; Belletini et al. 1998) and has bounded linear elastic fracture energy given by

$$\int_D 2\mu |\mathcal{E}\mathbf{u}^0(t)|^2 + \lambda |\operatorname{div} \mathbf{u}^0(t)|^2 dx + \mathcal{G}_c \mathcal{H}^2(J_{\mathbf{u}^0(t)}) \leq C,$$

for $0 \leq t \leq T$ where $J_{\mathbf{u}^0(t)}$ denotes the evolving fracture surface. The deformation-crack set pair $(\mathbf{u}^0(t), J_{\mathbf{u}^0(t)})$ records the brittle fracture evolution of the limit dynamics.

Arguments identical to Lipton (2016) show that away from sets where $|S(\mathbf{y}, \mathbf{x}; \mathbf{u}^{\epsilon_k})| > S_c$, the limit \mathbf{u}^0 satisfies the linear elastic wave equation. This is stated as follows: Fix $\delta > 0$ and for $\epsilon_k < \delta$ and $0 \leq t \leq T$ consider the open set $D' \subset D$ for which points \mathbf{x} in D' and \mathbf{y} for which $|\mathbf{y} - \mathbf{x}| < \epsilon_k$ satisfy,

$$|S(\mathbf{y}, \mathbf{x}; \mathbf{u}^{\epsilon_k}(t))| < S_c(\mathbf{y}, \mathbf{x}).$$

Then the limit evolution $\mathbf{u}^0(t, \mathbf{x})$ evolves elastodynamically on D' and is governed by the balance of linear momentum expressed by the Navier-Lamé equations on the domain $[0, T] \times D'$ given by

$$\mathbf{u}_{tt}^0(t) = \operatorname{div} \boldsymbol{\sigma}(t) + \mathbf{b}(t), \text{ on } [0, T] \times D',$$

where the stress tensor $\boldsymbol{\sigma}$ is given by

$$\boldsymbol{\sigma} = \lambda I_d \operatorname{Tr}(\mathcal{E} \mathbf{u}^0) + 2\mu \mathcal{E} \mathbf{u}^0,$$

where I_d is the identity on \mathbb{R}^d and $\operatorname{Tr}(\mathcal{E} \mathbf{u}^0)$ is the trace of the strain. Here the second derivative \mathbf{u}_{tt}^0 is the time derivative in the sense of distributions of \mathbf{u}_t^0 , and $\operatorname{div} \boldsymbol{\sigma}$ is the divergence of the stress tensor $\boldsymbol{\sigma}$ in the distributional sense. This shows that sharp fracture evolutions can be approached by spatially smooth evolutions in the limit of vanishing nonlocality.

Acknowledgements This material is based upon work supported by the US Army Research Laboratory and the US Army Research Office under contract/grant number W911NF1610456.

References

- L. Ambrosio, A. Coscia, G. Dal Maso, Fine properties of functions with bounded deformation. *Arch. Ration. Mech. Anal.* **139**, 201–238 (1997)
- G. Bellettini, A. Coscia, G. Dal Maso, Compactness and lower semicontinuity properties. *Mathematische Zeitschrift* **228**, 337–351 (1998)
- P. Diehl, R. Lipton, M. Schweitzer, Numerical verification of a bond-based softening peridynamic model for small displacements: deducing material parameters from classical linear theory. University of Bonn Technical report, Institut für Numerische Simulation (2016)
- F. Demengel and G. Demengel, *Functional spaces for the theory of elliptic partial differential equations*. Springer Verlag, London (2012)
- B.K. Driver, Analysis tools with applications. Lecture notes (2003), www.math.ucsd.edu/~bdriver/240-01-02/Lecture_Notes/anal.pdf
- E. Emmrich, R.B. Lehoucq, D. Puhst, Peridynamics: a nonlocal continuum theory, in *Meshfree Methods for Partial Differential Equations VI*, ed. by M. Griebel, M.A. Schweitzer. Lecture notes in computational science and engineering, vol. 89 (Springer, Berlin/Heidelberg, 2013), pp. 45–65
- P.K. Jha, R. Lipton, Finite element approximation of nonlinear nonlocal models. arXiv preprint arXiv:1710.07661 (2017a)
- P.K. Jha, R. Lipton, Numerical analysis of peridynamic models in Hölder space. arXiv preprint arXiv:1701.02818 (2017b)
- P.K. Jha, R. Lipton, Numerical convergence of nonlinear nonlocal continuum models to local elastodynamics. arXiv preprint arXiv:1707.00398 (2017c)
- R. Lipton, Dynamic brittle fracture as a small horizon limit of peridynamics. *J. Elast.* **117**(1), 21–50 (2014)
- R. Lipton, Cohesive dynamics and brittle fracture. *J. Elast.* **124**(2), 143–191 (2016)
- R. Lipton, S. Silling, R. Lehoucq, Complex fracture nucleation and evolution with nonlocal elastodynamics. arXiv preprint arXiv:1602.00247 (2016)
- T. Mengesha, Q. Du, Nonlocal constrained value problems for a linear peridynamic navier equation. *J. Elast.* **116**(1), 27–51 (2014)
- S.A. Silling, Reformulation of elasticity theory for discontinuities and long-range forces. *J. Mech. Phys. Solids* **48**(1), 175–209 (2000)
- S.A. Silling, R.B. Lehoucq, Convergence of peridynamics to classical elasticity theory. *J. Elast.* **93**(1), 13–37 (2008)



Finite Differences and Finite Elements in Nonlocal Fracture Modeling: A Priori Convergence Rates

44

Prashant K. Jha and Robert Lipton

Contents

Introduction	1458
Review of the Nonlocal Model	1460
Peridynamics Equation	1461
Finite Difference Approximation	1463
Time Discretization	1463
Stability of the Energy for the Semi-discrete Approximation	1477
Finite Element Approximation	1480
Projection of Function in FE Space	1481
Semi-discrete Approximation	1481
Central Difference Time Discretization	1482
Convergence of Approximation	1483
Stability Condition for Linearized Peridynamics	1488
Conclusion	1492
References	1493

Abstract

In this chapter we present a rigorous convergence analysis of finite difference and finite element approximation of nonlinear nonlocal models. In the previous chapter, we considered a differentiable version of the original bond-based model introduced in Silling (J Mech Phys Solids 48(1):175–209, 2000). There we showed, for a fixed horizon of nonlocal interaction ϵ , that well-posed

P. K. Jha

Department of Mathematics, Louisiana State University, Baton Rouge, LA, USA
e-mail: prashant.j16o@gmail.com; jha@math.lsu.edu

R. Lipton (✉)

Department of Mathematics, and Center for Computation and Technology, Louisiana State University, Baton Rouge, LA, USA
e-mail: lipton@math.lsu.edu; lipton@lsu.edu

formulations of the model can be developed over Hölder spaces and Sobolev spaces. In this chapter we apply these formulations to show a priori convergence for the discrete finite difference and finite element methods. We show that the error made using the forward Euler in time and a finite difference (i.e., piecewise constant) discretization in space with time step Δt and spatial discretization h is of the order of $O(\Delta t + h/\epsilon^2)$. For a central difference approximation in time and piecewise linear finite element approximation in space, the approximation error is of the order of $O(\Delta t + h^2/\epsilon^2)$. We point out these are the first such error estimates for nonlinear nonlocal fracture formulations and are reported in Jha and Lipton (2017b Numerical analysis of nonlocal fracture models models in holder space. arXiv preprint arXiv:1701.02818. To appear in SIAM Journal on Numerical Analysis 2018) and Jha and Lipton (2017a, Finite element approximation of nonlocal fracture models. arXiv preprint arXiv:1710.07661). We then go on to prove the stability of the semi-discrete approximation and show that the energy of the discrete approximation is bounded in terms of work done by the body force and initial energy put into the system. We look forward to improvements and development of a posteriori error estimation in the coming years.

Keywords

Peridynamic modeling · Finite differences · Finite elements · Stability · Convergence

Introduction

In this chapter we present a rigorous convergence analysis of finite difference and finite element approximation of nonlinear nonlocal fracture models. The model considered in this work is a differentiable version of the original peridynamic bond-based model introduced in Silling (2000) and analyzed in Lipton (2014, 2016). It is a bond-based model characterized by a nonlinear double-well potential. As discussed in the previous chapter, the nonlocal evolution converges to a sharp fracture evolution with bounded Griffith fracture energy as the length scale of nonlocality ϵ tends to zero. In this limit the displacement field satisfies the linear elastic wave equation off the fracture set.

We first consider the forward Euler time discretization and finite difference approximations in space with a uniform square mesh in 2-d and cubic mesh in 3-d. The mesh size is taken to be h and the time step is Δt . An a priori bound on the error is obtained for solutions in the Hölder space $C_0^{0,\gamma}(D; \mathbb{R}^d)$, where $\gamma \in (0, 1]$ is the Hölder exponent, D is the material domain, and $d = 2, 3$ is the dimension. The rate of convergence is shown to be no larger than $O(\Delta t + h^\gamma/\epsilon^2)$. We also show stability of the semi-discrete approximation. The semi-discrete evolution is shown to be uniformly bounded in time in terms of initial energy and the work done by body force. In this chapter we prove all results for the forward Euler in time

discretization, and we refer to Jha and Lipton (2017b) for the general single-step time discretization.

Next we consider central differences in time and a finite element discretization in space using triangular or tetrahedral meshes and conforming linear elements. Assuming $H_0^2(D; \mathbb{R}^d)$ solutions, we estimate the error and obtain a convergence rate of $O(\Delta t + h^2/\epsilon^2)$. We show that the semi-discrete evolution for the finite element scheme is also stable in time. We provide a stability analysis of the fully discrete problem, for the linearized peridynamic force. For this case we exhibit a CFL-like stability condition for the time step Δt .

The results presented here show that convergence requires $h^\nu < \epsilon^2$ for the finite difference case while $h^2 < \epsilon^2$ (or $h < \epsilon$) for the finite element case. The technical reason for the appearance of the factor $1/\epsilon^2$ in these convergence rates is that we are numerically approximating a nonlinear but Lipschitz continuous vector valued ODE. Here the vector space is the space of square integrable displacement fields, and the $1/\epsilon^2$ factor is proportional to the Lipschitz constant of the nonlocal nonlinear force acting on mean square integrable displacement fields. Our results requiring $h < \epsilon$ are consistent with the earlier work of Tian and Du (2014) for linear nonlocal forces and finite element approximations applied to equilibrium problems. We point out that the nonlocal nonlinear models treated here are identified with sharp fracture evolutions as $\epsilon \rightarrow 0$ (see Lipton 2014, 2016). However a convergence rate with respect to ϵ remains to be established. We discuss this aspect in the conclusions section.

There is a rapidly growing literature in peridynamic modeling and analysis (see, e.g., Emmrich et al. 2007; Du and Zhou 2011; Foster et al. 2011; Aksoylu and Parks 2011; Du et al. 2013a; Dayal 2017; Emmrich et al. 2013; Mengesha and Du 2013; Lipton 2014, 2016; Lipton et al. 2016; Emmrich and Puhst 2016; Du, Tao, and Tian 2017; Lipton et al. 2018; Aksoylu and Mengesha 2010; Mengesha and Du 2013, 2014; Aksoylu and Unlu 2014). In Macek and Silling (2007), Gerstle et al. (2007), Littlewood (2010), the finite element method is applied to the peridynamics formulation for the simulation of cracks. In Du et al. (2013b), the finite element approximation of linear peridynamic models for general quasistatic evolutions is analyzed. For linear elastic local models, the stability of the general Newmark time discretization is shown in Baker (1976), Grote and Schötzau (2009), and Karaa (2012). This behavior is shown to persist for elastic nonlocal models in Guan and Gunzburger (2015). In Chen and Gunzburger (2011), the finite element approximation with continuous and discontinuous elements is developed for nonlocal problems in one dimension. A numerical analysis of linear peridynamics models for a 1-d bar has been carried out in Weckner and Emmrich (2005) and Bobaru et al. (2009). In Tian and Du (2014) and Tian et al. (2016a,b), an asymptotically compatible approximation scheme is identified. In Askari et al. (2008), Silling et al. (2010), Ha and Bobaru (2011), Agwai et al. (2011), Bobaru and Hu (2012), and Zhang et al. (2016), crack prediction and crack branching phenomenon are analyzed through peridynamics. The list of references is by no means complete; additional references to the literature can be found in this handbook.

Review of the Nonlocal Model

We define the strain associated with the displacement field $\mathbf{u}(x)$ as

$$S(\mathbf{u}) = S(y, x; \mathbf{u}) = \frac{\mathbf{u}(y) - \mathbf{u}(x)}{|y - x|} \cdot \mathbf{e}_{y-x} \text{ and } \mathbf{e}_{y-x} = \frac{y - x}{|y - x|}. \tag{1}$$

We consider the following type of potential (see Figs. 1 and 2)

$$W^\epsilon(S, y - x) = \omega(x)\omega(y) \frac{J^\epsilon(|y - x|)}{\epsilon} f(|y - x|S^2), \tag{2}$$

where the function $f : \mathbb{R}^+ \rightarrow \mathbb{R}$ is positive, smooth, and concave and satisfies the following properties

$$\lim_{r \rightarrow 0^+} \frac{f(r)}{r} = f'(0), \quad \lim_{r \rightarrow \infty} f(r) = f_\infty < \infty. \tag{3}$$

Fig. 1 Two-point potential $W^\epsilon(S, y - x)$ as a function of strain S for fixed $y - x$

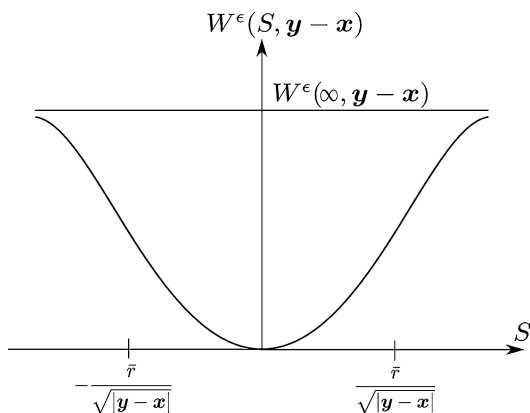
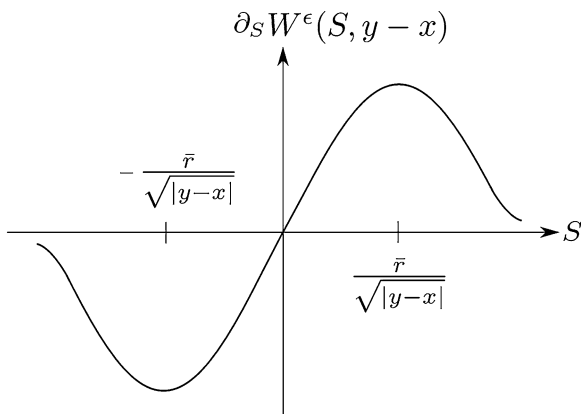


Fig. 2 Nonlocal force $\partial_S W^\epsilon(S, y - x)$ as a function of strain S for fixed $y - x$. Second derivative of $W^\epsilon(S, y - x)$ is zero at $\pm \bar{r}/\sqrt{|y - x|}$



The function $J^\epsilon(r) = J(r/\epsilon)$ is the influence function where $0 \leq J(|x|) \leq M$ for $x \in H_1(0)$ and $J = 0$ outside. The boundary function $0 \leq \omega(x) \leq 1$ takes the value 0 on the boundary ∂D of the material domain D . For $x \in D$, a distance ϵ away from boundary, $\omega(x)$ is 1 and smoothly decreases from 1 to zero as x approaches ∂D .

In the sequel we will set

$$\bar{\omega}(x) = \omega(x)\omega(x + \epsilon\xi) \tag{4}$$

and we assume

$$|\nabla\bar{\omega}| \leq C_{\omega_1} < \infty \quad \text{and} \quad |\nabla^2\bar{\omega}| \leq C_{\omega_2} < \infty.$$

The peridynamic force is written $-\nabla PD^\epsilon$ and given by

$$\begin{aligned} & -\nabla PD^\epsilon(\mathbf{u})(x) \\ &= \frac{2}{\epsilon^d \omega_d} \int_{H_\epsilon(x)} \partial_S W^\epsilon(S, y-x) \frac{y-x}{|y-x|} dy \\ &= \frac{4}{\epsilon^{d+1} \omega_d} \int_{H_\epsilon(x)} \omega(x)\omega(y) J^\epsilon(|y-x|) f'(|y-x| S(\mathbf{u})^2) S(\mathbf{u}) e_{y-x} dy, \end{aligned} \tag{5}$$

Peridynamics Equation

Let $u : [0, T] \times D \rightarrow \mathbb{R}^d$ be the displacement field such that it satisfies the following evolution equation

$$\rho \partial_{tt}^2 \mathbf{u}(t, x) = -\nabla PD^\epsilon(\mathbf{u}(t))(x) + b(t, x), \tag{6}$$

where $b(t, x)$ is the body force and ρ is the density. We will assume $\rho = 1$ throughout the chapter. The initial condition is given by

$$\mathbf{u}(0, x) = \mathbf{u}_0(x) \quad \text{and} \quad \partial_t \mathbf{u}(0, x) = \mathbf{v}_0(x) \tag{7}$$

and the boundary condition is given by

$$\mathbf{u}(t, x) = 0 \quad \forall x \in \partial D, \forall t \in [0, T]. \tag{8}$$

Throughout this chapter, we will assume $\mathbf{u} = 0$ on the boundary ∂D and is extended outside D by zero.

Additionally we can also write the evolution in weak form by multiplying Eq. 6 by a smooth test function $\tilde{\mathbf{u}}$ with $\tilde{\mathbf{u}} = 0$ on ∂D to get

$$(\ddot{\mathbf{u}}(t), \tilde{\mathbf{u}}) = (-\nabla PD^\epsilon(\mathbf{u}(t)), \tilde{\mathbf{u}}) + (b(t), \tilde{\mathbf{u}}).$$

We denote L^2 dot product of u, v as (u, v) . An integration by parts easily shows for all smooth u, v taking zero boundary values that

$$(-\nabla PD^\epsilon(\mathbf{u}), \mathbf{v}) = -a^\epsilon(\mathbf{u}, \mathbf{v}),$$

where

$$\begin{aligned} a^\epsilon(\mathbf{u}, \mathbf{v}) &= \frac{2}{\epsilon^{d+1}\omega_d} \int_D \int_{H_\epsilon(x)} \omega(x)\omega(y)J^\epsilon(|y-x|) \\ &\quad f'(|y-x|S(\mathbf{u})^2)|y-x|S(\mathbf{u})S(\mathbf{v})dydx. \end{aligned} \tag{9}$$

Weak form of the evolution in terms of operator a^ϵ becomes

$$(\ddot{\mathbf{u}}(t), \tilde{\mathbf{u}}) + a^\epsilon(\mathbf{u}(t), \tilde{\mathbf{u}}) = (b(t), \tilde{\mathbf{u}}). \tag{10}$$

Last we introduce the peridynamic energy. The total energy $\mathcal{E}^\epsilon(\mathbf{u})(t)$ is given by the sum of kinetic and potential energy given by

$$\mathcal{E}^\epsilon(\mathbf{u})(t) = \frac{1}{2} \|\dot{\mathbf{u}}(t)\|_{L^2(D;\mathbb{R}^d)}^2 + PD^\epsilon(\mathbf{u}(t)), \tag{11}$$

where potential energy PD^ϵ is given by

$$PD^\epsilon(\mathbf{u}) = \int_D \left[\frac{1}{\epsilon^d \omega_d} \int_{H_\epsilon(x)} W^\epsilon(S(\mathbf{u}), y-x) dy \right] dx.$$

We state the following equation which will be used later in the chapter

$$\frac{d}{dt} \mathcal{E}^\epsilon(\mathbf{u})(t) = (\ddot{\mathbf{u}}(t), \dot{\mathbf{u}}(t)) - (-\nabla PD^\epsilon(\mathbf{u}(t)), \dot{\mathbf{u}}(t)). \tag{12}$$

In order to develop the approximation theory in the following sections, we find it convenient to write the evolution Eq. 6 as an equivalent first order system with $y_1(t) = \mathbf{u}(t)$ and $y_2(t) = \mathbf{v}(t)$ with $\mathbf{v}(t) = \partial_t \mathbf{u}(t)$. Let $y = (y_1, y_2)^T$ where at each time y_1, y_2 belong to the same function space V , and let $F^\epsilon(y, t) = (F_1^\epsilon(y, t), F_2^\epsilon(y, t))^T$ such that

$$F_1^\epsilon(y, t) := y_2 \tag{13}$$

$$F_2^\epsilon(y, t) := -\nabla PD^\epsilon(y_1) + \mathbf{b}(t). \tag{14}$$

The initial boundary value associated with the evolution Eq. 6 is equivalent to the initial boundary value problem for the first order system given by

$$\frac{d}{dt}y = F^\epsilon(y, t), \tag{15}$$

with initial condition given by $y(0) = (\mathbf{u}_0, \mathbf{v}_0)^T \in X = V \times V$.

To establish the error estimates, we will use the Lipschitz property of the peridynamic force in $X = L^2(D; \mathbb{R}^d) \times L^2(D; \mathbb{R}^d)$. It is given by Theorem 6.1 of Lipton 2016.

Theorem 1.

$$\|F^\epsilon(y, t) - F^\epsilon(z, t)\|_X \leq \frac{L}{\epsilon^2} \|y - z\|_X \quad \forall y, z \in X, \forall t \in [0, T] \tag{16}$$

for all $y, z \in L^2(D; \mathbb{R}^d)^2$.

Here L does not depend on \mathbf{u}, \mathbf{v} .

Finite Difference Approximation

In this section, we present the finite difference scheme and compute the rate of convergence. We also consider the semi-discrete approximation and prove the bound on energy of semi-discrete evolution in terms of initial energy and the work done by body forces.

Let h be the size of a mesh and Δt be the size of time step. We will keep ϵ fixed and assume that $h < \epsilon < 1$. Let $D_h = D \cap (h\mathbb{Z})^d$ be the discretization of material domain. Let $i \in \mathbb{Z}^d$ be the index such that $\mathbf{x}_i = h\mathbf{i} \in D$. Let U_i be the unit cell of volume h^d corresponding to the grid point \mathbf{x}_i , see Fig. 3. The exact solution evaluated at grid points is denoted by $(\mathbf{u}_i(t), \mathbf{v}_i(t))$.

Time Discretization

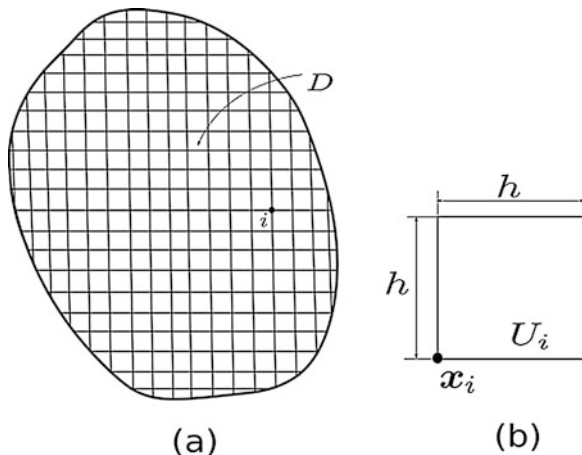
Let $[0, T] \cap (\Delta t\mathbb{Z})$ be the discretization of time domain where Δt is the size of time step. Denote fully discrete solution at $(t^k = k\Delta t, \mathbf{x}_i = ih)$ as $(\hat{\mathbf{u}}_i^k, \hat{\mathbf{v}}_i^k)$. Similarly, the exact solution evaluated at space-time grid points is denoted by $(\mathbf{u}_i^k, \mathbf{v}_i^k)$. We enforce the boundary condition $\hat{\mathbf{u}}_i^k = \mathbf{0}$ for all $\mathbf{x}_i \notin D$ and for all k .

We begin with the forward Euler time discretization, with respect to velocity, and the finite difference scheme for $(\hat{\mathbf{u}}_i^k, \hat{\mathbf{v}}_i^k)$ is written

$$\frac{\hat{\mathbf{u}}_i^{k+1} - \hat{\mathbf{u}}_i^k}{\Delta t} = \hat{\mathbf{v}}_i^{k+1} \tag{17}$$

$$\frac{\hat{\mathbf{v}}_i^{k+1} - \hat{\mathbf{v}}_i^k}{\Delta t} = -\nabla PD^\epsilon(\hat{\mathbf{u}}^k)(\mathbf{x}_i) + \mathbf{b}_i^k \tag{18}$$

Fig. 3 (a) Typical mesh of size h . (b) Unit cell U_i corresponding to material point \mathbf{x}_i



The scheme is complemented with the discretized initial conditions $\hat{\mathbf{u}}_i^0 = (\hat{\mathbf{u}}_0)_i$ and $\hat{\mathbf{v}}_i^0 = (\hat{\mathbf{v}}_0)_i$. If we substitute Eq. 17 into Eq. 18, we get the standard central difference scheme in time for second order in time differential equation. Here we have assumed, without loss of generality, $\rho = 1$.

The piecewise constant extensions of the discrete sets $\{\hat{\mathbf{u}}_i^k\}_{i \in \mathbb{Z}^d}$ and $\{\hat{\mathbf{v}}_i^k\}_{i \in \mathbb{Z}^d}$ are given by

$$\hat{\mathbf{u}}^k(\mathbf{x}) := \sum_{i, \mathbf{x}_i \in D} \hat{\mathbf{u}}_i^k \chi_{U_i}(\mathbf{x})$$

$$\hat{\mathbf{v}}^k(\mathbf{x}) := \sum_{i, \mathbf{x}_i \in D} \hat{\mathbf{v}}_i^k \chi_{U_i}(\mathbf{x})$$

In this way we represent the finite difference solution as a piecewise constant function. We will show that this function provides an L^2 approximation of the exact solution.

Convergence Results

In this section we provide upper bounds on the rate of convergence of the discrete approximation to the solution of the peridynamic evolution. The L^2 approximation error E^k at time t^k , for $0 < t^k \leq T$, is defined as

$$E^k := \left\| \hat{\mathbf{u}}^k - \mathbf{u}^k \right\|_{L^2(D; \mathbb{R}^d)} + \left\| \hat{\mathbf{v}}^k - \mathbf{v}^k \right\|_{L^2(D; \mathbb{R}^d)}$$

The upper bound on the convergence rate of the approximation error is given by the following theorem.

Theorem 2 (Convergence of finite difference approximation (forward Euler time discretization)). *Let $\epsilon > 0$ be fixed. Let (\mathbf{u}, \mathbf{v}) be the solution of peridynamic equation Eq. 15. We assume $\mathbf{u}, \mathbf{v} \in C^2([0, T]; C_0^{0,\gamma}(D; \mathbb{R}^d))$. Then the forward Euler time discretization, and finite difference spatial discretization scheme given by Eqs. 17 and 18, is consistent in both time and spatial discretization and converges to the exact solution uniformly in time with respect to the $L^2(D; \mathbb{R}^d)$ norm. If we assume the error at the initial step is zero, then the error E^k at time t^k is bounded and to leading order in the time step Δt satisfies*

$$\sup_{0 \leq k \leq T/\Delta t} E^k \leq O\left(C_t \Delta t + C_s \frac{h^\gamma}{\epsilon^2}\right), \tag{19}$$

where constants C_s and C_t are independent of h and Δt and C_s depends on the Hölder norm of the solution and C_t depends on the L^2 norms of time derivatives of the solution.

Here we have assumed the initial error to be zero for ease of exposition only.

We remark that the explicit constants leading to Eq. 19 can be large. The inequality that delivers Eq. 19 is given to leading order by

$$\sup_{0 \leq k \leq T/\Delta t} E^k \leq \exp\left[T(1 + 6\bar{C}/\epsilon^2)\right] T \left[C_t \Delta t + (C_s/\epsilon^2)h^\gamma\right], \tag{20}$$

where the constants \bar{C} , C_t , and C_s are given by Eqs. 43, 45, and 46. The explicit constant C_t depends on the spatial L^2 norm of the time derivatives of the solution, and C_s depends on the spatial Hölder continuity of the solution and the constant \bar{C} . This constant is bounded independently of horizon ϵ . Although the constants are necessarily pessimistic, they deliver a priori error estimates. These constants are discussed in Jha and Lipton (2017a) in the context of fracture experiments. Fracture experiments are on the order of hundreds of μ -sec, and the size of the constants in the estimate for finite element simulations remains small for tens of μ -sec. For finite element schemes, we have a priori estimates with constants that stay small for same order of magnitude in time as fracture experiments. These features are discussed in Jha and Lipton (2017a,b).

Error Analysis

We define the L^2 projections of the actual solutions onto the space of piecewise constant functions defined over the cells U_i . These are given as follows. Let $(\tilde{\mathbf{u}}_i^k, \tilde{\mathbf{v}}_i^k)$ be the average of the exact solution $(\mathbf{u}^k, \mathbf{v}^k)$ in the unit cell U_i given by

$$\begin{aligned} \tilde{\mathbf{u}}_i^k &:= \frac{1}{h^d} \int_{U_i} \mathbf{u}^k(\mathbf{x}) d\mathbf{x} \\ \tilde{\mathbf{v}}_i^k &:= \frac{1}{h^d} \int_{U_i} \mathbf{v}^k(\mathbf{x}) d\mathbf{x} \end{aligned}$$

and the L^2 projection of the solution onto piecewise constant functions is $(\tilde{\mathbf{u}}^k, \tilde{\mathbf{v}}^k)$ given by

$$\tilde{\mathbf{u}}^k(\mathbf{x}) := \sum_{i, \mathbf{x}_i \in D} \tilde{\mathbf{u}}_i^k \chi_{U_i}(\mathbf{x}) \tag{21}$$

$$\tilde{\mathbf{v}}^k(\mathbf{x}) := \sum_{i, \mathbf{x}_i \in D} \tilde{\mathbf{v}}_i^k \chi_{U_i}(\mathbf{x}) \tag{22}$$

The error between $(\hat{\mathbf{u}}^k, \hat{\mathbf{v}}^k)^T$ and $(\mathbf{u}(t^k), \mathbf{v}(t^k))^T$ is now split into two parts. From the triangle inequality, we have

$$\begin{aligned} \|\hat{\mathbf{u}}^k - \mathbf{u}(t^k)\|_{L^2(D; \mathbb{R}^d)} &\leq \|\hat{\mathbf{u}}^k - \tilde{\mathbf{u}}^k\|_{L^2(D; \mathbb{R}^d)} + \|\tilde{\mathbf{u}}^k - \mathbf{u}^k\|_{L^2(D; \mathbb{R}^d)} \\ \|\hat{\mathbf{v}}^k - \mathbf{v}(t^k)\|_{L^2(D; \mathbb{R}^d)} &\leq \|\hat{\mathbf{v}}^k - \tilde{\mathbf{v}}^k\|_{L^2(D; \mathbb{R}^d)} + \|\tilde{\mathbf{v}}^k - \mathbf{v}^k\|_{L^2(D; \mathbb{R}^d)} \end{aligned}$$

In section “[Error Analysis for Approximation of \$L^2\$ Projection of the Exact Solution](#)” we will show that the error between the L^2 projections of the actual solution and the discrete approximation decays according to

$$\sup_{0 \leq k \leq T/\Delta t} \left(\|\hat{\mathbf{u}}^k - \tilde{\mathbf{u}}^k\|_{L^2(D; \mathbb{R}^d)} + \|\hat{\mathbf{v}}^k - \tilde{\mathbf{v}}^k\|_{L^2(D; \mathbb{R}^d)} \right) = O\left(\Delta t + \frac{h^\gamma}{\epsilon^2}\right). \tag{23}$$

In what follows we can estimate the terms

$$\|\tilde{\mathbf{u}}^k - \mathbf{u}(t^k)\|_{L^2_0} \quad \text{and} \quad \|\tilde{\mathbf{v}}^k - \mathbf{v}(t^k)\|_{L^2_0} \tag{24}$$

and show they go to zero at a rate of h^γ uniformly in time. The estimates given by Eq. 23 together with the $O(h^\gamma)$ estimates for Eq. 24 establish Theorem 2. We now establish the L^2 estimates for the differences $\tilde{\mathbf{u}}^k - \mathbf{u}(t^k)$ and $\tilde{\mathbf{v}}^k - \mathbf{v}(t^k)$.

We write

$$\begin{aligned} &\|\tilde{\mathbf{u}}^k - \mathbf{u}^k\|_{L^2(D; \mathbb{R}^d)}^2 \\ &= \sum_{i, \mathbf{x}_i \in D} \int_{U_i} \left| \tilde{\mathbf{u}}^k(\mathbf{x}) - \mathbf{u}^k(\mathbf{x}) \right|^2 d\mathbf{x} \\ &= \sum_{i, \mathbf{x}_i \in D} \int_{U_i} \left| \frac{1}{h^d} \int_{U_i} (\mathbf{u}^k(\mathbf{y}) - \mathbf{u}^k(\mathbf{x})) d\mathbf{y} \right|^2 d\mathbf{x} \\ &= \sum_{i, \mathbf{x}_i \in D} \int_{U_i} \left[\frac{1}{h^{2d}} \int_{U_i} \int_{U_i} (\mathbf{u}^k(\mathbf{y}) - \mathbf{u}^k(\mathbf{x})) \cdot (\mathbf{u}^k(\mathbf{z}) - \mathbf{u}^k(\mathbf{x})) d\mathbf{y} d\mathbf{z} \right] d\mathbf{x} \end{aligned}$$

$$\leq \sum_{i, \mathbf{x}_i \in D} \int_{U_i} \left[\frac{1}{h^d} \int_{U_i} |\mathbf{u}^k(\mathbf{y}) - \mathbf{u}^k(\mathbf{x})|^2 d\mathbf{y} \right] d\mathbf{x} \tag{25}$$

where we used Cauchy’s inequality and Jensen’s inequality. For $\mathbf{x}, \mathbf{y} \in U_i$, $|\mathbf{x} - \mathbf{y}| \leq ch$, where $c = \sqrt{2}$ for $d = 2$ and $c = \sqrt{3}$ for $d = 3$. Since $\mathbf{u} \in C_0^{0,\gamma}$ we have

$$\begin{aligned} |\mathbf{u}^k(\mathbf{x}) - \mathbf{u}^k(\mathbf{y})| &= |\mathbf{x} - \mathbf{y}|^\gamma \frac{|\mathbf{u}^k(\mathbf{y}) - \mathbf{u}^k(\mathbf{x})|}{|\mathbf{x} - \mathbf{y}|^\gamma} \\ &\leq c^\gamma h^\gamma \|\mathbf{u}^k\|_{C^{0,\gamma}(D; \mathbb{R}^d)} \leq c^\gamma h^\gamma \sup_t \|\mathbf{u}(t)\|_{C^{0,\gamma}(D; \mathbb{R}^d)} \end{aligned} \tag{26}$$

and substitution in Eq. 25 gives

$$\begin{aligned} \|\tilde{\mathbf{u}}^k - \mathbf{u}^k\|_{L^2(D; \mathbb{R}^d)}^2 &\leq c^{2\gamma} h^{2\gamma} \sum_{i, \mathbf{x}_i \in D} \int_{U_i} d\mathbf{x} \left(\sup_t \|\mathbf{u}(t)\|_{C^{0,\gamma}(D; \mathbb{R}^d)} \right)^2 \\ &\leq c^{2\gamma} |D| h^{2\gamma} \left(\sup_t \|\mathbf{u}(t)\|_{C^{0,\gamma}(D; \mathbb{R}^d)} \right)^2. \end{aligned}$$

A similar estimate can be derived for $\|\tilde{\mathbf{v}}^k - \mathbf{v}^k\|_{L^2}$, and substitution of the estimates into Eq. 24 gives

$$\sup_k \left(\|\tilde{\mathbf{u}}^k - \mathbf{u}(t^k)\|_{L^2(D; \mathbb{R}^d)} + \|\tilde{\mathbf{v}}^k - \mathbf{v}(t^k)\|_{L^2(D; \mathbb{R}^d)} \right) = O(h^\gamma).$$

In the next section, we establish the error estimate (Eq. 23) for forward Euler in section “[Error Analysis for Approximation of \$L^2\$ Projection of the Exact Solution.](#)”

Error Analysis for Approximation of L^2 Projection of the Exact Solution

In this subsection, we estimate the difference between approximate solution $(\hat{\mathbf{u}}^k, \hat{\mathbf{v}}^k)$ and the L^2 projection of the exact solution onto piecewise constant functions given by $(\tilde{\mathbf{u}}^k, \tilde{\mathbf{v}}^k)$ (see Eqs. 21 and 22). Let the differences be denoted by $\mathbf{e}^k(\mathbf{u}) := \hat{\mathbf{u}}^k - \tilde{\mathbf{u}}^k$ and $\mathbf{e}^k(\mathbf{v}) := \hat{\mathbf{v}}^k - \tilde{\mathbf{v}}^k$, and their evaluations at grid points are $\mathbf{e}_i^k(\mathbf{u}) := \hat{\mathbf{u}}_i^k - \tilde{\mathbf{u}}_i^k$ and $\mathbf{e}_i^k(\mathbf{v}) := \hat{\mathbf{v}}_i^k - \tilde{\mathbf{v}}_i^k$. Subtracting $(\tilde{\mathbf{u}}_i^{k+1} - \tilde{\mathbf{u}}_i^k)/\Delta t$ from Eq. 17 gives

$$\begin{aligned} &\frac{\hat{\mathbf{u}}_i^{k+1} - \hat{\mathbf{u}}_i^k}{\Delta t} - \frac{\tilde{\mathbf{u}}_i^{k+1} - \tilde{\mathbf{u}}_i^k}{\Delta t} \\ &= \hat{\mathbf{v}}_i^{k+1} - \frac{\tilde{\mathbf{u}}_i^{k+1} - \tilde{\mathbf{u}}_i^k}{\Delta t} \\ &= \hat{\mathbf{v}}_i^{k+1} - \tilde{\mathbf{v}}_i^{k+1} + \left(\tilde{\mathbf{v}}_i^{k+1} - \frac{\partial \tilde{\mathbf{u}}_i^{k+1}}{\partial t} \right) + \left(\frac{\partial \tilde{\mathbf{u}}_i^{k+1}}{\partial t} - \frac{\tilde{\mathbf{u}}_i^{k+1} - \tilde{\mathbf{u}}_i^k}{\Delta t} \right). \end{aligned}$$

Taking the average over unit cell U_i of the exact peridynamic equation Eq. 15 at time t^k , we will get $\tilde{\mathbf{v}}_i^{k+1} - \frac{\partial \tilde{\mathbf{u}}_i^{k+1}}{\partial t} = 0$. Therefore, the equation for $\mathbf{e}_i^k(u)$ is given by

$$\mathbf{e}_i^{k+1}(u) = \mathbf{e}_i^k(u) + \Delta t \mathbf{e}_i^{k+1}(v) + \Delta t \tau_i^k(u), \tag{27}$$

where we identify the discretization error as

$$\tau_i^k(u) := \frac{\partial \tilde{\mathbf{u}}_i^{k+1}}{\partial t} - \frac{\tilde{\mathbf{u}}_i^{k+1} - \tilde{\mathbf{u}}_i^k}{\Delta t}. \tag{28}$$

Similarly, we subtract $(\tilde{\mathbf{v}}_i^{k+1} - \tilde{\mathbf{v}}_i^k)/\Delta t$ from Eq. 18 and add and subtract terms to get

$$\begin{aligned} \frac{\hat{\mathbf{v}}_i^{k+1} - \hat{\mathbf{v}}_i^k}{\Delta t} - \frac{\tilde{\mathbf{v}}_i^{k+1} - \tilde{\mathbf{v}}_i^k}{\Delta t} &= -\nabla PD^\epsilon(\hat{\mathbf{u}}^k)(\mathbf{x}_i) + \mathbf{b}_i^k - \frac{\partial \mathbf{v}_i^k}{\partial t} + \left(\frac{\partial \mathbf{v}_i^k}{\partial t} - \frac{\tilde{\mathbf{v}}_i^{k+1} - \tilde{\mathbf{v}}_i^k}{\Delta t} \right) \\ &= -\nabla PD^\epsilon(\hat{\mathbf{u}}^k)(\mathbf{x}_i) + \mathbf{b}_i^k - \frac{\partial \mathbf{v}_i^k}{\partial t} \\ &\quad + \left(\frac{\partial \tilde{\mathbf{v}}_i^k}{\partial t} - \frac{\tilde{\mathbf{v}}_i^{k+1} - \tilde{\mathbf{v}}_i^k}{\Delta t} \right) + \left(\frac{\partial \mathbf{v}_i^k}{\partial t} - \frac{\partial \tilde{\mathbf{v}}_i^k}{\partial t} \right), \end{aligned} \tag{29}$$

where we identify $\tau_i^k(v)$ as follows

$$\tau_i^k(v) := \frac{\partial \tilde{\mathbf{v}}_i^k}{\partial t} - \frac{\tilde{\mathbf{v}}_i^{k+1} - \tilde{\mathbf{v}}_i^k}{\Delta t}. \tag{30}$$

Note that in $\tau^k(u)$ we have $\frac{\partial \tilde{\mathbf{u}}_i^{k+1}}{\partial t}$, and from the exact peridynamic equation, we have

$$\mathbf{b}_i^k - \frac{\partial \mathbf{v}_i^k}{\partial t} = \nabla PD^\epsilon(\mathbf{u}^k)(\mathbf{x}_i). \tag{31}$$

Combining Eqs. 29, 30, and 31 gives

$$\begin{aligned} \mathbf{e}_i^{k+1}(v) &= \mathbf{e}_i^k(v) + \Delta t \tau_i^k(v) + \Delta t \left(\frac{\partial \mathbf{v}_i^k}{\partial t} - \frac{\partial \tilde{\mathbf{v}}_i^k}{\partial t} \right) \\ &\quad + \Delta t \left(-\nabla PD^\epsilon(\hat{\mathbf{u}}^k)(\mathbf{x}_i) + \nabla PD^\epsilon(\mathbf{u}^k)(\mathbf{x}_i) \right) \end{aligned}$$

$$\begin{aligned}
&= \mathbf{e}_i^k(v) + \Delta t \tau_i^k(v) + \Delta t \left(\frac{\partial \mathbf{v}_i^k}{\partial t} - \frac{\partial \tilde{\mathbf{v}}_i^k}{\partial t} \right) \\
&\quad + \Delta t \left(-\nabla PD^\epsilon(\hat{\mathbf{u}}^k)(\mathbf{x}_i) + \nabla PD^\epsilon(\bar{\mathbf{u}}^k)(\mathbf{x}_i) \right) \\
&\quad + \Delta t \left(-\nabla PD^\epsilon(\tilde{\mathbf{u}}^k)(\mathbf{x}_i) + \nabla PD^\epsilon(\mathbf{u}^k)(\mathbf{x}_i) \right).
\end{aligned}$$

The spatial discretization error $\sigma_i^k(u)$ and $\sigma_i^k(v)$ is given by

$$\sigma_i^k(u) := \left(-\nabla PD^\epsilon(\tilde{\mathbf{u}}^k)(\mathbf{x}_i) + \nabla PD^\epsilon(\mathbf{u}^k)(\mathbf{x}_i) \right) \quad (32)$$

$$\sigma_i^k(v) := \frac{\partial \mathbf{v}_i^k}{\partial t} - \frac{\partial \tilde{\mathbf{v}}_i^k}{\partial t}. \quad (33)$$

We finally have

$$\begin{aligned}
\mathbf{e}_i^{k+1}(v) &= \mathbf{e}_i^k(v) + \Delta t \left(\tau_i^k(v) + \sigma_i^k(u) + \sigma_i^k(v) \right) \\
&\quad + \Delta t \left(-\nabla PD^\epsilon(\hat{\mathbf{u}}^k)(\mathbf{x}_i) + \nabla PD^\epsilon(\bar{\mathbf{u}}^k)(\mathbf{x}_i) \right). \quad (34)
\end{aligned}$$

We now show the consistency and stability properties of the numerical scheme.

Consistency

We deal with the error in time discretization and the error in spatial discretization error separately. The time discretization error follows easily using the Taylor's series, while the spatial discretization error uses properties of the nonlinear peridynamic force.

Time discretization: We first estimate the time discretization error. A Taylor series expansion is used to estimate $\tau_i^k(u)$ as follows

$$\begin{aligned}
\tau_i^k(u) &= \frac{1}{h^d} \int_{U_i} \left(\frac{\partial \mathbf{u}^k(\mathbf{x})}{\partial t} - \frac{\mathbf{u}^{k+1}(\mathbf{x}) - \mathbf{u}^k(\mathbf{x})}{\Delta t} \right) d\mathbf{x} \\
&= \frac{1}{h^d} \int_{U_i} \left(-\frac{1}{2} \frac{\partial^2 \mathbf{u}^k(\mathbf{x})}{\partial t^2} \Delta t + O((\Delta t)^2) \right) d\mathbf{x}.
\end{aligned}$$

Computing the L^2 norm of $\tau_i^k(u)$ and using Jensen's inequality gives

$$\begin{aligned}
\|\tau^k(u)\|_{L^2(D; \mathbb{R}^d)} &\leq \frac{\Delta t}{2} \left\| \frac{\partial^2 \mathbf{u}^k}{\partial t^2} \right\|_{L^2(D; \mathbb{R}^d)} + O((\Delta t)^2) \\
&\leq \frac{\Delta t}{2} \sup_t \left\| \frac{\partial^2 \mathbf{u}(t)}{\partial t^2} \right\|_{L^2(D; \mathbb{R}^d)} + O((\Delta t)^2).
\end{aligned}$$

Similarly, we have

$$\|\tau^k(v)\|_{L^2(D;\mathbb{R}^d)} = \frac{\Delta t}{2} \sup_t \left\| \frac{\partial^2 \mathbf{v}(t)}{\partial t^2} \right\|_{L^2(D;\mathbb{R}^d)} + O((\Delta t)^2).$$

Spatial discretization: We now estimate the spatial discretization error. Substituting the definition of $\tilde{\mathbf{v}}^k$ and following the similar steps employed in Eq. 26 gives

$$\begin{aligned} |\sigma_i^k(v)| &= \left| \frac{\partial v_i^k}{\partial t} - \frac{1}{h^d} \int_{U_i} \frac{\partial \mathbf{v}^k(\mathbf{x})}{\partial t} d\mathbf{x} \right| \leq c^\gamma h^\gamma \int_{U_i} \frac{1}{|\mathbf{x}_i - \mathbf{x}|^\gamma} \left| \frac{\partial \mathbf{v}^k(\mathbf{x}_i)}{\partial t} - \frac{\partial \mathbf{v}^k(\mathbf{x})}{\partial t} \right| d\mathbf{x} \\ &\leq c^\gamma h^\gamma \left\| \frac{\partial \mathbf{v}^k}{\partial t} \right\|_{C^{0,\gamma}(D;\mathbb{R}^d)} \leq c^\gamma h^\gamma \sup_t \left\| \frac{\partial \mathbf{v}(t)}{\partial t} \right\|_{C^{0,\gamma}(D;\mathbb{R}^d)}. \end{aligned}$$

Taking the L^2 norm of error $\sigma_i^k(v)$ and substituting the estimate above delivers

$$\|\sigma^k(v)\|_{L^2(D;\mathbb{R}^d)} \leq h^\gamma c^\gamma \sqrt{|D|} \sup_t \left\| \frac{\partial \mathbf{v}(t)}{\partial t} \right\|_{C^{0,\gamma}(D;\mathbb{R}^d)}.$$

Now we estimate $|\sigma_i^k(u)|$. Note that the force $-\nabla PD^\epsilon(\mathbf{u})(\mathbf{x})$ can be written as follows

$$\begin{aligned} -\nabla PD^\epsilon(\mathbf{u})(\mathbf{x}) &= \frac{4}{\epsilon^{d+1} \omega_d} \int_{H_\epsilon(\mathbf{x})} \omega(\mathbf{x}) \omega(\mathbf{y}) J\left(\frac{|\mathbf{y} - \mathbf{x}|}{\epsilon}\right) f'(|\mathbf{y} - \mathbf{x}| S(\mathbf{y}, \mathbf{x}; \mathbf{u})^2) \\ &S(\mathbf{y}, \mathbf{x}; \mathbf{u}) \frac{\mathbf{y} - \mathbf{x}}{|\mathbf{y} - \mathbf{x}|} d\mathbf{y} = \frac{4}{\epsilon \omega_d} \int_{H_1(\mathbf{0})} \omega(\mathbf{x}) \omega(\mathbf{x} + \epsilon \boldsymbol{\xi}) J(|\boldsymbol{\xi}|) f'(\epsilon |\boldsymbol{\xi}| \\ &S(\mathbf{x} + \epsilon \boldsymbol{\xi}, \mathbf{x}; \mathbf{u})^2) S(\mathbf{x} + \epsilon \boldsymbol{\xi}, \mathbf{x}; \mathbf{u}) \frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|} d\boldsymbol{\xi}. \end{aligned}$$

where we substituted $\partial_S W^\epsilon$ using Eq. 2. In the second step, we introduced the change in variable $\mathbf{y} = \mathbf{x} + \epsilon \boldsymbol{\xi}$.

Let $F_1 : \mathbb{R} \rightarrow \mathbb{R}$ be defined as $F_1(S) = f(S^2)$. Then $F_1'(S) = f'(S^2)2S$. Using the definition of F_1 , we have

$$2Sf'(\epsilon |\boldsymbol{\xi}| S^2) = \frac{F_1'(\sqrt{\epsilon |\boldsymbol{\xi}|} S)}{\sqrt{\epsilon |\boldsymbol{\xi}|}}.$$

Because f is assumed to be positive, smooth, and concave and is bounded far away, we have the following bound on derivatives of F_1

$$\sup_r |F_1'(r)| = F_1'(\bar{r}) =: C_1 \tag{35}$$

$$\sup_r |F_1''(r)| = \max\{F_1''(0), F_1''(\hat{u})\} =: C_2 \tag{36}$$

$$\sup_r |F_1'''(r)| = \max\{F_1'''(\tilde{u}_2), F_1'''(\tilde{u}_2)\} =: C_3. \tag{37}$$

where \bar{r} is the inflection point of $f(r^2)$, i.e., $F_1''(\bar{r}) = 0$. $\{0, \hat{u}\}$ are the maxima of $F_1''(r)$. $\{\tilde{u}, \tilde{u}\}$ are the maxima of $F_1'''(r)$. By chain rule and by considering the assumption on f , we can show that $\bar{r}, \hat{u}, \tilde{u}_2, \tilde{u}_2$ exist and the C_1, C_2, C_3 are bounded. Figures 4, 5, and 6 show the generic graphs of $F_1'(r), F_1''(r)$, and $F_1'''(r)$, respectively.

Fig. 4 Generic plot of $F_1'(r)$. $|F_1'(r)|$ is bounded by $|F_1'(\bar{r})|$

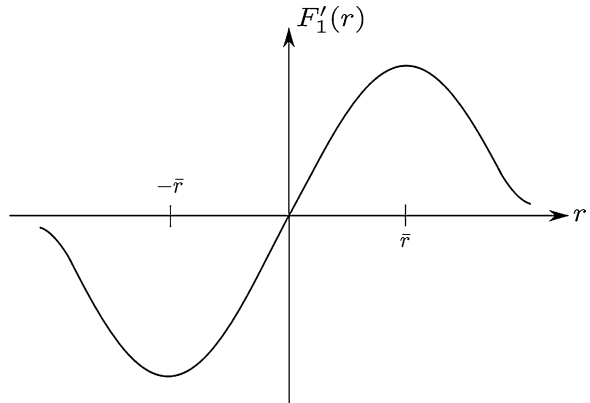


Fig. 5 Generic plot of $F_1''(r)$. At $\pm\bar{r}$, $F_1''(r) = 0$. At $\pm\hat{u}$, $F_1'''(r) = 0$

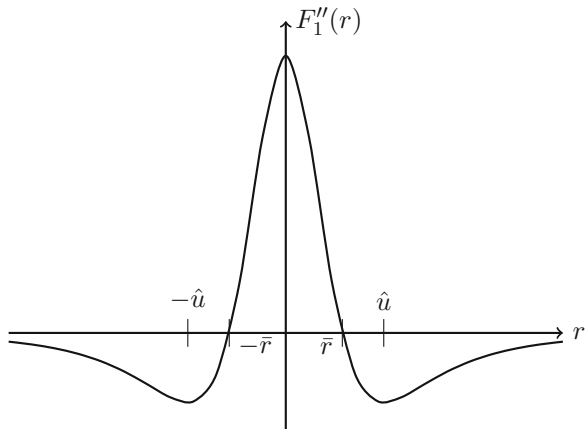
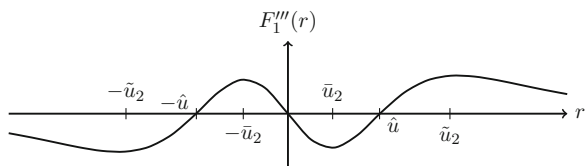


Fig. 6 Generic plot of $F_1'''(r)$. At $\pm\tilde{u}_2$ and $\pm\tilde{u}_2$, $F_1'''' = 0$



The nonlocal force $-\nabla PD^\epsilon$ can be written as

$$\begin{aligned}
 &-\nabla PD^\epsilon(\mathbf{u})(\mathbf{x}) \\
 &= \frac{2}{\epsilon\omega_d} \int_{H_1(\mathbf{0})} \omega(\mathbf{x})\omega(\mathbf{x} + \epsilon\xi)J(|\xi|)F'_1(\sqrt{\epsilon|\xi|}S(\mathbf{x} + \epsilon\xi, \mathbf{x}; \mathbf{u})) \frac{1}{\sqrt{\epsilon|\xi|}} \frac{\xi}{|\xi|} d\xi.
 \end{aligned} \tag{38}$$

To simplify the calculations, we use the following notation

$$\begin{aligned}
 \bar{\mathbf{u}}(\mathbf{x}) &:= \mathbf{u}(\mathbf{x} + \epsilon\xi) - \mathbf{u}(\mathbf{x}), \\
 \bar{\mathbf{v}}(\mathbf{y}) &:= \mathbf{v}(\mathbf{y} + \epsilon\xi) - \mathbf{v}(\mathbf{y}), \\
 (\mathbf{u} - \mathbf{v})(\mathbf{x}) &:= \mathbf{u}(\mathbf{x}) - \mathbf{v}(\mathbf{x}),
 \end{aligned}$$

and $\overline{(\mathbf{u} - \mathbf{v})}(\mathbf{x})$ is defined similar to $\bar{\mathbf{u}}(\mathbf{x})$. Also, let

$$s = \epsilon|\xi|, \quad \mathbf{e} = \frac{\xi}{|\xi|}.$$

In what follows, we will come across the integral of type $\int_{H_1(\mathbf{0})} J(|\xi|) |\xi|^{-\alpha} d\xi$. Recall that $0 \leq J(|\xi|) \leq M$ for all $\xi \in H_1(\mathbf{0})$ and $J(|\xi|) = 0$ for $\xi \notin H_1(\mathbf{0})$. Therefore, let

$$\bar{J}_\alpha := \frac{1}{\omega_d} \int_{H_1(\mathbf{0})} J(|\xi|) |\xi|^{-\alpha} d\xi. \tag{39}$$

With notations above, we note that $S(\mathbf{x} + \epsilon\xi, \mathbf{x}; \mathbf{u}) = \bar{\mathbf{u}}(\mathbf{x}) \cdot \mathbf{e}/s$. $-\nabla PD^\epsilon$ can be written as

$$\begin{aligned}
 -\nabla PD^\epsilon(\mathbf{u})(\mathbf{x}) &= \frac{2}{\epsilon\omega_d} \int_{H_1(\mathbf{0})} \omega(\mathbf{x})\omega(\mathbf{x} + \epsilon\xi)J(|\xi|)F'_1(\bar{\mathbf{u}}(\mathbf{x}) \cdot \mathbf{e}/\sqrt{s}) \frac{1}{\sqrt{s}} \mathbf{e} d\xi.
 \end{aligned} \tag{40}$$

We estimate $|-\nabla PD^\epsilon(\mathbf{u})(\mathbf{x}) - (-\nabla PD^\epsilon(\mathbf{v})(\mathbf{x}))|$.

$$\begin{aligned}
 &|-\nabla PD^\epsilon(\mathbf{u})(\mathbf{x}) - (-\nabla PD^\epsilon(\mathbf{v})(\mathbf{x}))| \\
 &\leq \left| \frac{2}{\epsilon\omega_d} \int_{H_1(\mathbf{0})} \omega(\mathbf{x})\omega(\mathbf{x} + \epsilon\xi)J(|\xi|) \frac{(F'_1(\bar{\mathbf{u}}(\mathbf{x}) \cdot \mathbf{e}/\sqrt{s}) - F'_1(\bar{\mathbf{v}}(\mathbf{x}) \cdot \mathbf{e}/\sqrt{s}))}{\sqrt{s}} \mathbf{e} d\xi \right| \\
 &\leq \left| \frac{2}{\epsilon\omega_d} \int_{H_1(\mathbf{0})} J(|\xi|) \frac{1}{\sqrt{s}} |F'_1(\bar{\mathbf{u}}(\mathbf{x}) \cdot \mathbf{e}/\sqrt{s}) - F'_1(\bar{\mathbf{v}}(\mathbf{x}) \cdot \mathbf{e}/\sqrt{s})| d\xi \right| \\
 &\leq \sup_r |F''_1(r)| \left| \frac{2}{\epsilon\omega_d} \int_{H_1(\mathbf{0})} J(|\xi|) \frac{1}{\sqrt{s}} |\bar{\mathbf{u}}(\mathbf{x}) \cdot \mathbf{e}/\sqrt{s} - \bar{\mathbf{v}}(\mathbf{x}) \cdot \mathbf{e}/\sqrt{s}| d\xi \right|
 \end{aligned}$$

$$\leq \frac{2C_2}{\epsilon\omega_d} \left| \int_{H_1(\mathbf{0})} J(|\xi|) \frac{|\bar{\mathbf{u}}(\mathbf{x}) - \bar{\mathbf{v}}(\mathbf{x})|}{\epsilon|\xi|} d\xi \right|. \quad (41)$$

Here we have used the fact that $|\omega(\mathbf{x})| \leq 1$ and for a vector \mathbf{e} such that $|\mathbf{e}| = 1$, $|\mathbf{a} \cdot \mathbf{e}| \leq |\mathbf{a}|$ holds and $|\alpha\mathbf{e}| \leq |\alpha|$ holds for all $\mathbf{a} \in \mathbb{R}^d$, $\alpha \in \mathbb{R}$.

We use the notation $\bar{\mathbf{u}}^k(\mathbf{x}) := \mathbf{u}^k(\mathbf{x} + \epsilon\xi) - \mathbf{u}^k(\mathbf{x})$ and $\bar{\tilde{\mathbf{u}}}^k(\mathbf{x}) := \tilde{\mathbf{u}}(\mathbf{x} + \epsilon\xi) - \tilde{\mathbf{u}}^k(\mathbf{x})$ and choose $\mathbf{u} = \mathbf{u}^k$ and $\mathbf{v} = \tilde{\mathbf{u}}^k$ in Eq. 41 to find that

$$\begin{aligned} |\sigma_i^k(u)| &= \left| -\nabla PD^\epsilon(\bar{\tilde{\mathbf{u}}}^k)(\mathbf{x}_i) + \nabla PD^\epsilon(\mathbf{u}^k)(\mathbf{x}_i) \right| \\ &\leq \frac{2C_2}{\epsilon\omega_d} \left| \int_{H_1(\mathbf{0})} J(|\xi|) \frac{|\mathbf{u}^k(\mathbf{x}_i + \epsilon\xi) - \tilde{\mathbf{u}}^k(\mathbf{x}_i + \epsilon\xi) - (\mathbf{u}^k(\mathbf{x}_i) - \tilde{\mathbf{u}}^k(\mathbf{x}_i))|}{\epsilon|\xi|} d\xi \right|. \end{aligned} \quad (42)$$

Here C_2 is the maximum of the second derivative of the profile describing the potential given by Eq. 36. Following the earlier analysis (see Eq. 26), we find that

$$\begin{aligned} \left| \mathbf{u}^k(\mathbf{x}_i + \epsilon\xi) - \tilde{\mathbf{u}}^k(\mathbf{x}_i + \epsilon\xi) \right| &\leq c^\gamma h^\gamma \sup_t \|\mathbf{u}(t)\|_{C^{0,\gamma}(D;\mathbb{R}^d)} \\ \left| \mathbf{u}^k(\mathbf{x}_i) - \tilde{\mathbf{u}}^k(\mathbf{x}_i) \right| &\leq c^\gamma h^\gamma \sup_t \|\mathbf{u}(t)\|_{C^{0,\gamma}(D;\mathbb{R}^d)}. \end{aligned}$$

For reference, we define the constant

$$\bar{C} = \frac{C_2}{\omega_d} \int_{H_1(\mathbf{0})} J(|\xi|) \frac{1}{|\xi|} d\xi. \quad (43)$$

We now focus on Eq. 42. We substitute the above two inequalities to get

$$\begin{aligned} |\sigma_i^k(u)| &\leq \frac{2C_2}{\epsilon^2\omega_d} \left| \int_{H_1(\mathbf{0})} J(|\xi|) \frac{1}{|\xi|} \right. \\ &\quad \left. \left(\left| \mathbf{u}^k(\mathbf{x}_i + \epsilon\xi) - \tilde{\mathbf{u}}^k(\mathbf{x}_i + \epsilon\xi) \right| + \left| \mathbf{u}^k(\mathbf{x}_i) - \tilde{\mathbf{u}}^k(\mathbf{x}_i) \right| \right) d\xi \right| \\ &\leq 4h^\gamma c^\gamma \frac{\bar{C}}{\epsilon^2} \sup_t \|\mathbf{u}(t)\|_{C^{0,\gamma}(D;\mathbb{R}^d)}. \end{aligned}$$

Therefore, we have

$$\|\sigma^k(u)\|_{L^2(D;\mathbb{R}^d)} \leq h^\gamma \left(4c^\gamma \sqrt{|D|} \frac{\bar{C}}{\epsilon^2} \sup_t \|\mathbf{u}(t)\|_{C^{0,\gamma}(D;\mathbb{R}^d)} \right).$$

This completes the proof of consistency of numerical approximation.

Stability

Let e^k be the total error at the k th time step. It is defined as

$$e^k := \|e^k(u)\|_{L^2(D;\mathbb{R}^d)} + \|e^k(v)\|_{L^2(D;\mathbb{R}^d)}.$$

To simplify the calculations, we define new term τ as

$$\begin{aligned} \tau := & \sup_t \left(\|\tau^k(u)\|_{L^2(D;\mathbb{R}^d)} + \|\tau^k(v)\|_{L^2(D;\mathbb{R}^d)} \right. \\ & \left. + \|\sigma^k(u)\|_{L^2(D;\mathbb{R}^d)} + \|\sigma^k(v)\|_{L^2(D;\mathbb{R}^d)} \right). \end{aligned}$$

From our consistency analysis, we know that to leading order

$$\tau \leq C_t \Delta t + \frac{C_s}{\epsilon^2} h^\gamma \tag{44}$$

where,

$$C_t := \frac{1}{2} \sup_t \left\| \frac{\partial^2 \mathbf{u}(t)}{\partial t^2} \right\|_{L^2(D;\mathbb{R}^d)} + \frac{1}{2} \sup_t \left\| \frac{\partial^3 \mathbf{u}(t)}{\partial t^3} \right\|_{L^2(D;\mathbb{R}^d)}, \tag{45}$$

$$C_s := c^\gamma \sqrt{|D|} \left[\epsilon^2 \sup_t \left\| \frac{\partial^2 \mathbf{u}(t)}{\partial t^2} \right\|_{C^{0,\gamma}(D;\mathbb{R}^d)} + 4\bar{C} \sup_t \|\mathbf{u}(t)\|_{C^{0,\gamma}(D;\mathbb{R}^d)} \right]. \tag{46}$$

We take L^2 norm of Eqs. 27 and 34 and add them. Noting the definition of τ as above, we get

$$\begin{aligned} e^{k+1} \leq & e^k + \Delta t \|e^{k+1}(v)\|_{L^2(D;\mathbb{R}^d)} + \Delta t \tau \\ & + \Delta t \left(\sum_i h^d \left| -\nabla PD^\epsilon(\hat{\mathbf{u}}^k)(x_i) + \nabla PD^\epsilon(\tilde{\mathbf{u}}^k)(x_i) \right|^2 \right)^{1/2}. \end{aligned} \tag{47}$$

We only need to estimate the last term in the above equation. Similar to the Eq. 42, we have

$$\begin{aligned} & \left| -\nabla PD^\epsilon(\hat{\mathbf{u}}^k)(x_i) + \nabla PD^\epsilon(\tilde{\mathbf{u}}^k)(x_i) \right| \\ & \leq \frac{2C_2}{\epsilon^2 \omega_d} \left| \int_{H_1(\mathbf{0})} J(|\xi|) \frac{1}{|\xi|} \left| \hat{\mathbf{u}}^k(x_i + \epsilon \xi) - \tilde{\mathbf{u}}^k(x_i + \epsilon \xi) - (\hat{\mathbf{u}}^k(x_i) - \tilde{\mathbf{u}}^k(x_i)) \right| d\xi \right| \\ & = \frac{2C_2}{\epsilon^2 \omega_d} \left| \int_{H_1(\mathbf{0})} J(|\xi|) \frac{1}{|\xi|} |e^k(u)(x_i + \epsilon \xi) - e^k(u)(x_i)| d\xi \right| \\ & \leq \frac{2C_2}{\epsilon^2 \omega_d} \left| \int_{H_1(\mathbf{0})} J(|\xi|) \frac{1}{|\xi|} (|e^k(u)(x_i + \epsilon \xi)| + |e^k(u)(x_i)|) d\xi \right|. \end{aligned}$$

By $\mathbf{e}^k(u)(\mathbf{x})$ we mean evaluation of piecewise extension of set $\{\mathbf{e}_i^k(u)\}_i$ at \mathbf{x} . We proceed further as follows

$$\begin{aligned} & \left| -\nabla PD^\epsilon(\hat{\mathbf{u}}^k)(\mathbf{x}_i) + \nabla PD^\epsilon(\tilde{\mathbf{u}}^k)(\mathbf{x}_i) \right|^2 \\ & \leq \left(\frac{2C_2}{\epsilon^2 \omega_d} \right)^2 \int_{H_1(\mathbf{0})} \int_{H_1(\mathbf{0})} J(|\xi|) J(|\eta|) \frac{1}{|\xi|} \frac{1}{|\eta|} \\ & \quad (|\mathbf{e}^k(u)(\mathbf{x}_i + \epsilon \xi)| + |\mathbf{e}^k(u)(\mathbf{x}_i)|) (|\mathbf{e}^k(u)(\mathbf{x}_i + \epsilon \eta)| + |\mathbf{e}^k(u)(\mathbf{x}_i)|) d\xi d\eta. \end{aligned}$$

Using inequality $|ab| \leq (|a|^2 + |b|^2)/2$, we get

$$\begin{aligned} & (|\mathbf{e}^k(u)(\mathbf{x}_i + \epsilon \xi)| + |\mathbf{e}^k(u)(\mathbf{x}_i)|) (|\mathbf{e}^k(u)(\mathbf{x}_i + \epsilon \eta)| + |\mathbf{e}^k(u)(\mathbf{x}_i)|) \\ & \leq 3 \left(|\mathbf{e}^k(u)(\mathbf{x}_i + \epsilon \xi)|^2 + |\mathbf{e}^k(u)(\mathbf{x}_i + \epsilon \eta)|^2 + |\mathbf{e}^k(u)(\mathbf{x}_i)|^2 \right), \end{aligned}$$

and

$$\begin{aligned} & \sum_{\mathbf{x}_i \in D} h^d \left| -\nabla PD^\epsilon(\hat{\mathbf{u}}^k)(\mathbf{x}_i) + \nabla PD^\epsilon(\tilde{\mathbf{u}}^k)(\mathbf{x}_i) \right|^2 \\ & \leq \left(\frac{2C_2}{\epsilon^2 \omega_d} \right)^2 \int_{H_1(\mathbf{0})} \int_{H_1(\mathbf{0})} J(|\xi|) J(|\eta|) \frac{1}{|\xi|} \frac{1}{|\eta|} \\ & \quad \sum_{\mathbf{x}_i \in D} h^d 3 \left(|\mathbf{e}^k(u)(\mathbf{x}_i + \epsilon \xi)|^2 + |\mathbf{e}^k(u)(\mathbf{x}_i + \epsilon \eta)|^2 + |\mathbf{e}^k(u)(\mathbf{x}_i)|^2 \right) d\xi d\eta. \end{aligned}$$

Since $\mathbf{e}^k(u)(\mathbf{x}) = \sum_{\mathbf{x}_i \in D} \mathbf{e}_i^k(u) \chi_{U_i}(\mathbf{x})$, we have

$$\sum_{\mathbf{x}_i \in D} h^d \left| -\nabla PD^\epsilon(\hat{\mathbf{u}}^k)(\mathbf{x}_i) + \nabla PD^\epsilon(\tilde{\mathbf{u}}^k)(\mathbf{x}_i) \right|^2 \leq \frac{(6\bar{C})^2}{\epsilon^4} \|\mathbf{e}^k(u)\|_{L^2(D; \mathbb{R}^d)}^2. \quad (48)$$

where \bar{C} is given by Eq. 43. In summary Eq. 48 shows the Lipschitz continuity of the peridynamic force with respect to the L^2 norm (see Eq. 16) expressed in this context as

$$\|\nabla PD^\epsilon(\hat{\mathbf{u}}^k)(\mathbf{x}) - \nabla PD^\epsilon(\tilde{\mathbf{u}}^k)\|_{L^2(D; \mathbb{R}^d)} \leq \frac{(6\bar{C})}{\epsilon^2} \|\mathbf{e}^k(u)\|_{L^2(D; \mathbb{R}^d)}. \quad (49)$$

Finally, we substitute above inequality in Eq. 47 to get

$$e^{k+1} \leq e^k + \Delta t \|\mathbf{e}^{k+1}(v)\|_{L^2(D; \mathbb{R}^d)} + \Delta t \tau + \Delta t \frac{6\bar{C}}{\epsilon^2} \|\mathbf{e}^k(u)\|_{L^2(D; \mathbb{R}^d)}$$

We add positive quantity $\Delta t \|e^{k+1}(u)\|_{L^2(D;\mathbb{R}^d)} + \Delta t 6\bar{C}/\epsilon^2 \|e^k(v)\|_{L^2(D;\mathbb{R}^d)}$ to the right side of above equation to get

$$\begin{aligned}
 e^{k+1} &\leq (1 + \Delta t 6\bar{C}/\epsilon^2)e^k + \Delta t e^{k+1} + \Delta t \tau \\
 \Rightarrow e^{k+1} &\leq \frac{(1 + \Delta t 6\bar{C}/\epsilon^2)}{1 - \Delta t} e^k + \frac{\Delta t}{1 - \Delta t} \tau.
 \end{aligned}$$

We recursively substitute e^j on above as follows

$$\begin{aligned}
 e^{k+1} &\leq \frac{(1 + \Delta t 6\bar{C}/\epsilon^2)}{1 - \Delta t} e^k + \frac{\Delta t}{1 - \Delta t} \tau \\
 &\leq \left(\frac{(1 + \Delta t 6\bar{C}/\epsilon^2)}{1 - \Delta t}\right)^2 e^{k-1} + \frac{\Delta t}{1 - \Delta t} \tau \left(1 + \frac{(1 + \Delta t 6\bar{C}/\epsilon^2)}{1 - \Delta t}\right) \\
 &\leq \dots \\
 &\leq \left(\frac{(1 + \Delta t 6\bar{C}/\epsilon^2)}{1 - \Delta t}\right)^{k+1} e^0 + \frac{\Delta t}{1 - \Delta t} \tau \sum_{j=0}^k \left(\frac{(1 + \Delta t 6\bar{C}/\epsilon^2)}{1 - \Delta t}\right)^{k-j}. \tag{50}
 \end{aligned}$$

Since $1/(1 - \Delta t) = 1 + \Delta t + \Delta t^2 + O(\Delta t^3)$, we have

$$\frac{(1 + \Delta t 6\bar{C}/\epsilon^2)}{1 - \Delta t} \leq 1 + (1 + 6\bar{C}/\epsilon^2)\Delta t + (1 + 6\bar{C}/\epsilon^2)\Delta t^2 + O(\bar{C}/\epsilon^2)O(\Delta t^3).$$

Now, for any $k \leq T/\Delta t$, using identity $(1 + a)^k \leq \exp[ka]$ for $a \leq 0$, we have

$$\begin{aligned}
 &\left(\frac{1 + \Delta t 6\bar{C}/\epsilon^2}{1 - \Delta t}\right)^k \\
 &\leq \exp[k(1 + 6\bar{C}/\epsilon^2)\Delta t + k(1 + 6\bar{C}/\epsilon^2)\Delta t^2 + kO(\bar{C}/\epsilon^2)O(\Delta t^3)] \\
 &\leq \exp[T(1 + 6\bar{C}/\epsilon^2) + T(1 + 6\bar{C}/\epsilon^2)\Delta t + O(T\bar{C}/\epsilon^2)O(\Delta t^2)].
 \end{aligned}$$

We write above equation in more compact form as follows

$$\begin{aligned}
 &\left(\frac{1 + \Delta t 6\bar{C}/\epsilon^2}{1 - \Delta t}\right)^k \\
 &\leq \exp[T(1 + 6\bar{C}/\epsilon^2)(1 + \Delta t + O(\Delta t^2))].
 \end{aligned}$$

We use above estimate in Eq. 50 and get the following inequality for e^k

$$\begin{aligned} e^{k+1} &\leq \exp [T(1 + 6\bar{C}/\epsilon^2)(1 + \Delta t + O(\Delta t^2))] (e^0 + (k + 1)\tau\Delta t/(1 - \Delta t)) \\ &\leq \exp [T(1 + 6\bar{C}/\epsilon^2)(1 + \Delta t + O(\Delta t^2))] (e^0 + T\tau(1 + \Delta t + O(\Delta t^2))). \end{aligned}$$

where we used the fact that $1/(1 - \Delta t) = 1 + \Delta t + O(\Delta t^2)$.

Assuming the error in initial data is zero, i.e., $e^0 = 0$, and noting the estimate of τ in Eq. 44, we have

$$\sup_k e^k \leq \exp [T(1 + 6\bar{C}/\epsilon^2)] T\tau$$

and we conclude to leading order that

$$\sup_k e^k \leq \exp [T(1 + 6\bar{C}/\epsilon^2)] T [C_t \Delta t + (C_s/\epsilon^2)h^\gamma], \quad (51)$$

Here the constants C_t and C_s are given by Eqs. 45 and 46. This shows the stability of the numerical scheme and Theorem 2 is proved.

Stability of the Energy for the Semi-discrete Approximation

We first spatially discretize the peridynamics equation (Eq. 6). Let $\{\hat{\mathbf{u}}_i(t)\}_{i, \mathbf{x}_i \in D}$ denote the semi-discrete approximate solution which satisfies the following, for all $t \in [0, T]$ and i such that $\mathbf{x}_i \in D$,

$$\ddot{\hat{\mathbf{u}}}_i(t) = -\nabla P D^\epsilon(\hat{\mathbf{u}}(t))(\mathbf{x}_i) + \mathbf{b}_i(t) \quad (52)$$

where $\hat{\mathbf{u}}(t)$ is the piecewise constant extension of discrete set $\{\hat{\mathbf{u}}_i(t)\}_i$ and is defined as

$$\hat{\mathbf{u}}(t, \mathbf{x}) := \sum_{i, \mathbf{x}_i \in D} \hat{\mathbf{u}}_i(t) \chi_{U_i}(\mathbf{x}). \quad (53)$$

The scheme is complemented with the discretized initial conditions $\hat{\mathbf{u}}_i(0) = \mathbf{u}_0(\mathbf{x}_i)$ and $\dot{\hat{\mathbf{u}}}_i(0) = \mathbf{v}_0(\mathbf{x}_i)$. We apply boundary condition by setting $\hat{\mathbf{u}}_i(t) = \mathbf{0}$ for all t and for all $\mathbf{x}_i \notin D$.

We have the stability of semi-discrete evolution.

Theorem 3 (Energy stability of the semi-discrete approximation). *Let $\{\hat{\mathbf{u}}_i(t)\}_i$ satisfy Eq. 52 and $\hat{\mathbf{u}}(t)$ is its piecewise constant extension. Similarly let $\hat{\mathbf{b}}(t, \mathbf{x})$ denote the piecewise constant extension of $\{\mathbf{b}(t, \mathbf{x}_i)\}_{i, \mathbf{x}_i \in D}$. Then the peridynamic energy \mathcal{E}^ϵ as defined in Eq. 11 satisfies, $\forall t \in [0, T]$,*

$$\mathcal{E}^\epsilon(\hat{\mathbf{u}})(t) \leq \left(\sqrt{\mathcal{E}^\epsilon(\hat{\mathbf{u}})(0)} + \frac{TC}{\epsilon^{3/2}} + \int_0^T \|\hat{\mathbf{b}}(s)\|_{L^2(D; \mathbb{R}^d)} ds \right)^2. \quad (54)$$

The constant C , defined in Eq. 59, is independent of ϵ and h .

Proof. We multiply Eq. 52 by $\chi_{U_i}(\mathbf{x})$ and sum over i and use definition of piecewise constant extension in Eq. 53 to get

$$\begin{aligned} \ddot{\hat{\mathbf{u}}}(t, \mathbf{x}) &= -\nabla P\hat{D}^\epsilon(\hat{\mathbf{u}}(t))(\mathbf{x}) + \hat{\mathbf{b}}(t, \mathbf{x}) \\ &= -\nabla PD^\epsilon(\hat{\mathbf{u}}(t))(\mathbf{x}) + \hat{\mathbf{b}}(t, \mathbf{x}) \\ &\quad + (-\nabla P\hat{D}^\epsilon(\hat{\mathbf{u}}(t))(\mathbf{x}) + \nabla PD^\epsilon(\hat{\mathbf{u}}(t))(\mathbf{x})) \end{aligned}$$

where $-\nabla P\hat{D}^\epsilon(\hat{\mathbf{u}}(t))(\mathbf{x})$ and $\hat{\mathbf{b}}(t, \mathbf{x})$ are given by

$$\begin{aligned} -\nabla P\hat{D}^\epsilon(\hat{\mathbf{u}}(t))(\mathbf{x}) &= \sum_{i, \mathbf{x}_i \in D} (-\nabla PD^\epsilon(\hat{\mathbf{u}}(t))(\mathbf{x}_i))\chi_{U_i}(\mathbf{x}) \\ \hat{\mathbf{b}}(t, \mathbf{x}) &= \sum_{i, \mathbf{x}_i \in D} \mathbf{b}(t, \mathbf{x}_i)\chi_{U_i}(\mathbf{x}). \end{aligned}$$

We define set as follows

$$\sigma(t, \mathbf{x}) := -\nabla P\hat{D}^\epsilon(\hat{\mathbf{u}}(t))(\mathbf{x}) + \nabla PD^\epsilon(\hat{\mathbf{u}}(t))(\mathbf{x}). \tag{55}$$

We use the following result which we will show after few steps

$$\|\sigma(t)\|_{L^2(D; \mathbb{R}^d)} \leq \frac{C}{\epsilon^{3/2}}. \tag{56}$$

We then have

$$\ddot{\hat{\mathbf{u}}}(t, \mathbf{x}) = -\nabla PD^\epsilon(\hat{\mathbf{u}}(t))(\mathbf{x}) + \hat{\mathbf{b}}(t, \mathbf{x}) + \sigma(t, \mathbf{x}). \tag{57}$$

Multiply above with $\dot{\hat{\mathbf{u}}}(t)$ and integrate over D to get

$$\begin{aligned} (\ddot{\hat{\mathbf{u}}}(t), \dot{\hat{\mathbf{u}}}(t)) &= (-\nabla PD^\epsilon(\hat{\mathbf{u}}(t)), \dot{\hat{\mathbf{u}}}(t)) \\ &\quad + (\hat{\mathbf{b}}(t), \dot{\hat{\mathbf{u}}}(t)) + (\sigma(t), \dot{\hat{\mathbf{u}}}(t)). \end{aligned}$$

Consider energy $\mathcal{E}^\epsilon(\hat{\mathbf{u}})(t)$ given by Eq. 11 and note the identity Eq. 12, to have

$$\begin{aligned} \frac{d}{dt} \mathcal{E}^\epsilon(\hat{\mathbf{u}})(t) &= (\hat{\mathbf{b}}(t), \dot{\hat{\mathbf{u}}}(t)) + (\sigma(t), \dot{\hat{\mathbf{u}}}(t)) \\ &\leq \left(\|\hat{\mathbf{b}}(t)\|_{L^2(D; \mathbb{R}^d)} + \|\sigma(t)\|_{L^2(D; \mathbb{R}^d)} \right) \|\dot{\hat{\mathbf{u}}}(t)\|_{L^2(D; \mathbb{R}^d)}, \end{aligned}$$

where we used Hölder inequality in last step. Since $PD^\epsilon(\mathbf{u})$ is positive for any \mathbf{u} , we have

$$\|\dot{\hat{\mathbf{u}}}(t)\| \leq 2\sqrt{\frac{1}{2}\|\dot{\hat{\mathbf{u}}}(t)\|_{L^2(D;\mathbb{R}^d)}^2 + PD^\epsilon(\hat{\mathbf{u}}(t))} = 2\sqrt{\mathcal{E}^\epsilon(\hat{\mathbf{u}})(t)}.$$

Using above, we get

$$\frac{1}{2} \frac{d}{dt} \mathcal{E}^\epsilon(\hat{\mathbf{u}})(t) \leq \left(\|\hat{\mathbf{b}}(t)\|_{L^2(D;\mathbb{R}^d)} + \|\sigma(t)\|_{L^2(D;\mathbb{R}^d)} \right) \sqrt{\mathcal{E}^\epsilon(\hat{\mathbf{u}})(t)}.$$

Let $\delta > 0$ be some arbitrary but fixed real number and let $A(t) = \delta + \mathcal{E}^\epsilon(\hat{\mathbf{u}})(t)$. Then

$$\frac{1}{2} \frac{d}{dt} A(t) \leq \left(\|\hat{\mathbf{b}}(t)\|_{L^2(D;\mathbb{R}^d)} + \|\sigma(t)\|_{L^2(D;\mathbb{R}^d)} \right) \sqrt{A(t)}.$$

Using the fact that $\frac{1}{\sqrt{A(t)}} \frac{d}{dt} A(t) = 2 \frac{d}{dt} \sqrt{A(t)}$, we have

$$\begin{aligned} \sqrt{A(t)} &\leq \sqrt{A(0)} + \int_0^t \left(\|\hat{\mathbf{b}}(s)\|_{L^2(D;\mathbb{R}^d)} + \|\sigma(s)\|_{L^2(D;\mathbb{R}^d)} \right) ds \\ &\leq \sqrt{A(0)} + \frac{TC}{\epsilon^{3/2}} + \int_0^T \|\hat{\mathbf{b}}(s)\|_{L^2(D;\mathbb{R}^d)} ds. \end{aligned}$$

where we used bound on $\|\sigma(s)\|_{L^2(D;\mathbb{R}^d)}$ from Eq. 56. Noting that $\delta > 0$ is arbitrary, we send it to zero to get

$$\sqrt{\mathcal{E}^\epsilon(\hat{\mathbf{u}})(t)} \leq \sqrt{\mathcal{E}^\epsilon(\hat{\mathbf{u}})(0)} + \frac{TC}{\epsilon^{3/2}} + \int_0^T \|\hat{\mathbf{b}}(s)\| ds,$$

and Eq. 54 follows by taking square of above equation.

It remains to show in Eq. 56. To simplify the calculations, we use the following notations: let $\xi \in H_1(\mathbf{0})$ and let

$$\begin{aligned} s_\xi &= \epsilon|\xi|, e_\xi = \frac{\xi}{|\xi|}, \bar{\omega}(\mathbf{x}) = \omega(\mathbf{x})\omega(\mathbf{x} + \epsilon\xi), \\ S_\xi(\mathbf{x}) &= \frac{\hat{\mathbf{u}}(t, \mathbf{x} + \epsilon\xi) - \hat{\mathbf{u}}(t, \mathbf{x})}{s_\xi} \cdot e_\xi. \end{aligned}$$

With above notations and using expression of $-\nabla PD^\epsilon$ from Eq. 38, we have for $\mathbf{x} \in U_i$

$$\begin{aligned} |\sigma(t, \mathbf{x})| &= |-\nabla PD^\epsilon(\hat{\mathbf{u}}(t))(x_i) + \nabla PD^\epsilon(\hat{\mathbf{u}}(t))(x)| \\ &= \left| \frac{2}{\epsilon\omega_d} \int_{H_1(\mathbf{0})} \frac{J(|\xi|)}{\sqrt{s_\xi}} (\bar{\omega}(x_i) F'_1(\sqrt{s_\xi} S_\xi(x_i)) - \bar{\omega}(x) F'_1(\sqrt{s_\xi} S_\xi(x))) e_\xi d\xi \right| \end{aligned}$$

$$\begin{aligned} &\leq \frac{2}{\epsilon\omega_d} \int_{H_1(\mathbf{0})} \frac{J(|\xi|)}{\sqrt{s_\xi}} |\bar{\omega}(\mathbf{x}_i)F'_1(\sqrt{s_\xi}S_\xi(\mathbf{x}_i)) - \bar{\omega}(\mathbf{x})F'_1(\sqrt{s_\xi}S_\xi(\mathbf{x}))| d\xi \\ &\leq \frac{2}{\epsilon\omega_d} \int_{H_1(\mathbf{0})} \frac{J(|\xi|)}{\sqrt{s_\xi}} (|\bar{\omega}(\mathbf{x}_i)F'_1(\sqrt{s_\xi}S_\xi(\mathbf{x}_i))| + |\bar{\omega}(\mathbf{x})F'_1(\sqrt{s_\xi}S_\xi(\mathbf{x}))|) d\xi. \end{aligned} \tag{58}$$

Using the fact that $0 \leq \omega(\mathbf{x}) \leq 1$ and $|F'_1(r)| \leq C_1$, where C_1 is $\sup_r |F'_1(r)|$, we get

$$|\sigma(t, \mathbf{x})| \leq \frac{4C_1\bar{J}_{1/2}}{\epsilon^{3/2}}.$$

where $\bar{J}_{1/2} = (1/\omega_d) \int_{H_1(\mathbf{0})} J(|\xi|)|\xi|^{-1/2}d\xi$.

Taking the L^2 norm of $\sigma(t, \mathbf{x})$, we get

$$\|\sigma(t)\|_{L^2(D;\mathbb{R}^d)}^2 = \sum_{i,\mathbf{x}_i \in D} \int_{U_i} |\sigma(t, \mathbf{x})|^2 d\mathbf{x} \leq \left(\frac{4C_1\bar{J}_{1/2}}{\epsilon^{3/2}}\right)^2 \sum_{i,\mathbf{x}_i \in D} \int_{U_i} d\mathbf{x}$$

thus

$$\|\sigma(t)\|_{L^2(D;\mathbb{R}^d)} \leq \frac{4C_1\bar{J}_{1/2}\sqrt{|D|}}{\epsilon^{3/2}} = \frac{C}{\epsilon^{3/2}}$$

where

$$C := 4C_1\bar{J}_{1/2}\sqrt{|D|}. \tag{59}$$

This completes the proof.

Finite Element Approximation

Let V_h be the approximation of $H_0^2(D, \mathbb{R}^d)$ associated with linear continuous interpolation associated with the mesh \mathcal{T}_h (triangular or tetrahedral) where h denotes the size of finite element mesh. Let $\mathcal{I}_h(u)$ be defined as below

$$\mathcal{I}_h(u)(x) = \sum_{T \in \mathcal{T}_h} \left[\sum_{i \in N_T} u(x_i)\phi_i(x) \right]$$

where N_T is the set of global indices of nodes associated to finite element T , ϕ_i is the linear interpolation function associated to node i , and x_i is the material coordinate of node i .

Assuming that the size of each element in the mesh \mathcal{T}_h is bounded by h , we have (see, e.g., [Theorem 4.6, Arnold 2011])

$$\|u - \mathcal{I}_h(u)\| \leq ch^2 \|u\|_2, \quad \forall u \in H_0^2(D; \mathbb{R}^d). \quad (60)$$

Projection of Function in FE Space

Let $r_h(u) \in V_h$ be the projection of $u \in H_0^2(D; \mathbb{R}^d)$. It is defined as

$$\|u - r_h(u)\| = \inf_{\tilde{u} \in V_h} \|u - \tilde{u}\|. \quad (61)$$

It also satisfies the following

$$(r_h(u), \tilde{u}) = (u, \tilde{u}), \quad \forall \tilde{u} \in V_h. \quad (62)$$

Since $\mathcal{I}_h(u) \in V_h$, we get an upper bound on right-hand side term and we have

$$\|u - r_h(u)\| \leq ch^2 \|u\|_2 \quad \forall u \in H_0^2(D; \mathbb{R}^d). \quad (63)$$

Semi-discrete Approximation

Let $u_h(t) \in V_h$ be the approximation of $u(t)$ which satisfies the following

$$(\ddot{u}_h, \tilde{u}) + a^\epsilon(u_h(t), \tilde{u}) = (b(t), \tilde{u}), \quad \forall \tilde{u} \in V_h. \quad (64)$$

We now show that the semi-discrete approximation is stable, i.e., energy at time t is bounded by initial energy and work done by the body force.

Theorem 4 (Stability of the semi-discrete approximation). *The semi-discrete scheme is energetically stable and the energy $\mathcal{E}^\epsilon(u_h)(t)$, defined in (11), satisfies the following bound*

$$\mathcal{E}^\epsilon(u_h)(t) \leq \left[\sqrt{\mathcal{E}^\epsilon(u_h)(0)} + \int_0^t \|b(\tau)\| d\tau \right]^2.$$

We note that, while proving the stability of semi-discrete scheme corresponding to nonlinear peridynamics, we do not require any assumption on strain $S(y, x; u_h)$.

Proof. Letting $\tilde{u} = \dot{u}_h(t)$ in (10) and noting the identity (12), we get

$$\frac{d}{dt} \mathcal{E}^\epsilon(u_h)(t) = (b(t), \dot{u}_h(t)) \leq \|b(t)\| \|\dot{u}_h(t)\|$$

We also have

$$\|\dot{u}_h(t)\| \leq 2\sqrt{\frac{1}{2}\|\dot{u}_h\|^2 + PD^\epsilon(u_h(t))} = 2\sqrt{\mathcal{E}^\epsilon(u_h)(t)}$$

where we used the fact that $PD^\epsilon(u)(t)$ is nonnegative. We have

$$\frac{d}{dt}\mathcal{E}^\epsilon(u_h)(t) \leq 2\sqrt{\mathcal{E}^\epsilon(u_h)(t)}\|b(t)\|.$$

Fix $\delta > 0$ and let $A(t) = \mathcal{E}^\epsilon(u_h(t)) + \delta$. Then, from above equation, we easily have

$$\frac{d}{dt}A(t) \leq 2\sqrt{A(t)}\|b(t)\| \Rightarrow \frac{1}{2}\frac{d}{dt}\sqrt{A(t)} \leq \|b(t)\|.$$

Noting that $\frac{1}{\sqrt{a(t)}}\frac{da(t)}{dt} = 2\frac{d}{dt}\sqrt{a(t)}$, integrating from $t = 0$ to τ , and relabeling τ as t , we get

$$\sqrt{A(t)} \leq \sqrt{A(0)} + \int_0^t \|b(s)\| ds.$$

Letting $\delta \rightarrow 0$ and taking the square of both sides proves the claim.

Central Difference Time Discretization

For illustration, we consider the central difference scheme and present the convergence rate for the central difference scheme for the fully nonlinear problem. We remark that the extension of these results to the general Newmark scheme is straightforward. We then consider a linearized peridynamics and demonstrate CFL-like conditions for stability of the fully discrete scheme.

Let Δt be the time step. The exact solution at $t^k = k\Delta t$ (or time step k) is denoted as (u^k, v^k) , with $v^k = \partial u^k / \partial t$, and the projection onto V_h at t^k is given by $(r_h(u^k), r_h(v^k))$. The solution of the discrete problem at time step k is denoted as (u_h^k, v_h^k) .

We approximate the initial data on displacement u_0 and velocity v_0 by their projection $r_h(u_0)$ and $r_h(v_0)$. Let $u_h^0 = r_h(u_0)$ and $v_h^0 = r_h(v_0)$. For $k \geq 1$, (u_h^k, v_h^k) satisfies, for all $\tilde{u} \in V_h$,

$$\begin{aligned} \left(\frac{u_h^{k+1} - u_h^k}{\Delta t}, \tilde{u}\right) &= (v_h^{k+1}, \tilde{u}), \\ \left(\frac{v_h^{k+1} - v_h^k}{\Delta t}, \tilde{u}\right) &= (-\nabla PD^\epsilon(u_h^k), \tilde{u}) + (b_h^k, \tilde{u}), \end{aligned} \tag{65}$$

where we denote projection of $b(t^k)$, $r_h(b(t^k))$, as b_h^k . Combining the two equations delivers central difference equation for u_h^k . We have

$$\left(\frac{u_h^{k+1} - 2u_h^k + u_h^{k-1}}{\Delta t^2}, \tilde{u} \right) = (-\nabla P D^\epsilon(u_h^k), \tilde{u}) + (b_h^k, \tilde{u}), \quad \forall \tilde{u} \in V_h. \quad (66)$$

For $k = 0$, we have $\forall \tilde{u} \in V_h$

$$\left(\frac{u_h^1 - u_h^0}{\Delta t^2}, \tilde{u} \right) = \frac{1}{2}(-\nabla P D^\epsilon(u_h^0), \tilde{u}) + \frac{1}{\Delta t}(v_h^0, \tilde{u}) + \frac{1}{2}(b_h^0, \tilde{u}). \quad (67)$$

We now show that finite element discretization converges to the exact solution.

Convergence of Approximation

In this section, we establish uniform bound on the discretization error and prove that approximate solution converges to the exact solution at the rate $C_l \Delta t + C_s h^2 / \epsilon^2$ for fixed $\epsilon > 0$. We first compare the exact solution with its projection in V_h and then compare the projection with approximate solution. We further divide the calculation of error between projection and approximate solution in two parts, namely, consistency analysis and error analysis.

Error E^k is given by

$$E^k := \|u_h^k - u(t^k)\| + \|v_h^k - v(t^k)\|.$$

We split the error as follows

$$E^k \leq (\|u^k - r_h(u^k)\| + \|v^k - r_h(v^k)\|) + (\|r_h(u^k) - u_h^k\| + \|r_h(v^k) - v_h^k\|),$$

where first term is error between exact solution and projections and second term is error between projections and approximate solution. Let

$$e_h^k(u) := r_h(u^k) - u_h^k \quad \text{and} \quad e_h^k(v) := r_h(v^k) - v_h^k \quad (68)$$

and

$$e^k := \|e_h^k(u)\| + \|e_h^k(v)\|. \quad (69)$$

Using (63), we have

$$E^k \leq C_p h^2 + e^k, \quad (70)$$

where

$$C_p := c \left[\sup_t \|u(t)\|_2 + \sup_t \left\| \frac{\partial u(t)}{\partial t} \right\|_2 \right]. \tag{71}$$

We have the following main result.

Theorem 5 (Convergence of the central difference approximation). *Let (u, v) be the exact solution of peridynamics equation in (6). Let (u_h^k, v_h^k) be the central difference approximation in time and piecewise linear finite element approximation in space solution of (65). If $u, v \in C^2([0, T], H_0^2(D; \mathbb{R}^d))$, then the scheme is consistent, and the error E^k satisfies the following bound*

$$\begin{aligned} & \sup_{k \leq T/\Delta t} E^k \\ &= C_p h^2 + \exp[T(1 + L/\epsilon^2)(1 + \Delta t + O(\Delta t^2))] \\ & \quad \left[e^0 + T(1 + \Delta t + O(\Delta t^2)) \left(C_t \Delta t + C_s \frac{h^2}{\epsilon^2} \right) \right] \end{aligned} \tag{72}$$

where constants $C_p, C_t,$ and C_s are given by (71) and (79). The constant L/ϵ^2 is the Lipschitz constant of $-\nabla P D^\epsilon(u)$ in L^2 (see Theorem 1).

If the error in initial data is zero, then E^k is of the order of $C_t \Delta t + C_s h^2/\epsilon^2$.

Error Analysis

We derive the equation for evolution of $e_h^k(u)$ as follows

$$\begin{aligned} & \left(\frac{u_h^{k+1} - u_h^k}{\Delta t} - \frac{r_h(u^{k+1}) - r_h(u^k)}{\Delta t}, \tilde{u} \right) \\ &= (v_h^{k+1}, \tilde{u}) - \left(\frac{r_h(u^{k+1}) - r_h(u^k)}{\Delta t}, \tilde{u} \right) \\ &= (v_h^{k+1}, \tilde{u}) - (r_h(v^{k+1}), \tilde{u}) + (r_h(v^{k+1}), \tilde{u}) - (v^{k+1}, \tilde{u}) \\ & \quad + (v^{k+1}, \tilde{u}) - \left(\frac{\partial u^{k+1}}{\partial t}, \tilde{u} \right) \\ & \quad + \left(\frac{\partial u^{k+1}}{\partial t}, \tilde{u} \right) - \left(\frac{u^{k+1} - u^k}{\Delta t}, \tilde{u} \right) \\ & \quad + \left(\frac{u^{k+1} - u^k}{\Delta t}, \tilde{u} \right) - \left(\frac{r_h(u^{k+1}) - r_h(u^k)}{\Delta t}, \tilde{u} \right). \end{aligned}$$

Using property $(r_h(u), \tilde{u}) = (u, \tilde{u})$ for $\tilde{u} \in V_h$ and the fact that $\frac{\partial u(t^{k+1})}{\partial t} = v^{k+1}$ where u is the exact solution, we get

$$\left(\frac{e_h^{k+1}(u) - e_h^k(u)}{\Delta t}, \tilde{u} \right) = (e_h^{k+1}(v), \tilde{u}) + \left(\frac{\partial u^{k+1}}{\partial t}, \tilde{u} \right) - \left(\frac{u^{k+1} - u^k}{\Delta t}, \tilde{u} \right). \quad (73)$$

Let $(\tau_h^k(u), \tau_h^k(v))$ be the consistency error in time discretization given by

$$\begin{aligned} \tau_h^k(u) &:= \frac{\partial u^{k+1}}{\partial t} - \frac{u^{k+1} - u^k}{\Delta t}, \\ \tau_h^k(v) &:= \frac{\partial v^k}{\partial t} - \frac{v^{k+1} - v^k}{\Delta t}. \end{aligned}$$

With above notation, we have

$$(e_h^{k+1}(u), \tilde{u}) = (e_h^k(u), \tilde{u}) + \Delta t (e_h^{k+1}(v), \tilde{u}) + \Delta t (\tau_h^k(u), \tilde{u}). \quad (74)$$

We now derive the equation for $e_h^k(v)$ as follows

$$\begin{aligned} &\left(\frac{v_h^{k+1} - v_h^k}{\Delta t} - \frac{r_h(v^{k+1}) - r_h(v^k)}{\Delta t}, \tilde{u} \right) \\ &= (-\nabla PD^\epsilon(u_h^k), \tilde{u}) + (b_h^k, \tilde{u}) - \left(\frac{r_h(v^{k+1}) - r_h(v^k)}{\Delta t}, \tilde{u} \right) \\ &= (-\nabla PD^\epsilon(u_h^k), \tilde{u}) + (b^k, \tilde{u}) - \left(\frac{\partial v^k}{\partial t}, \tilde{u} \right) \\ &\quad + \left(\frac{\partial v^k}{\partial t}, \tilde{u} \right) - \left(\frac{v^{k+1} - v^k}{\Delta t}, \tilde{u} \right) \\ &\quad + \left(\frac{v^{k+1} - v^k}{\Delta t}, \tilde{u} \right) - \left(\frac{r_h(v^{k+1}) - r_h(v^k)}{\Delta t}, \tilde{u} \right) \\ &= (-\nabla PD^\epsilon(u_h^k) + \nabla PD^\epsilon(u^k), \tilde{u}) + (b_h^k - b(t^k), \tilde{u}) \\ &\quad + \left(\frac{\partial v^k}{\partial t}, \tilde{u} \right) - \left(\frac{v^{k+1} - v^k}{\Delta t}, \tilde{u} \right) + \left(\frac{v^{k+1} - v^k}{\Delta t}, \tilde{u} \right) - \left(\frac{r_h(v^{k+1}) - r_h(v^k)}{\Delta t}, \tilde{u} \right) \\ &= (-\nabla PD^\epsilon(u_h^k) + \nabla PD^\epsilon(u^k), \tilde{u}) + \left(\frac{\partial v^k}{\partial t} - \frac{v^{k+1} - v^k}{\Delta t}, \tilde{u} \right) \end{aligned}$$

where we used the property of $r_h(u)$ and the fact that

$$(-\nabla PD^\epsilon(u^k), \tilde{u}) + (b^k, \tilde{u}) - \left(\frac{\partial v^k}{\partial t}, \tilde{u} \right) = 0, \quad \forall \tilde{u} \in H_0^2(D; \mathbb{R}^d).$$

We further divide the error in peridynamics force as follows

$$\begin{aligned} & (-\nabla PD^\epsilon(u_h^k) + \nabla PD^\epsilon(u^k), \tilde{u}) \\ &= (-\nabla PD^\epsilon(u_h^k) + \nabla PD^\epsilon(r_h(u^k)), \tilde{u}) + (-\nabla PD^\epsilon(r_h(u^k)) + \nabla PD^\epsilon(u^k), \tilde{u}). \end{aligned}$$

We will see in the next section that the second term is related to consistency error in spatial discretization. Therefore, we define another consistency error term $\sigma_{per,h}^k(u)$ as follows

$$\sigma_{per,h}^k(u) := -\nabla PD^\epsilon(r_h(u^k)) + \nabla PD^\epsilon(u^k). \tag{75}$$

After substituting the notations related to consistency errors, we get

$$\begin{aligned} (e_h^{k+1}(v), \tilde{u}) &= (e_h^k(v), \tilde{u}) + \Delta t (-\nabla PD^\epsilon(u_h^k) + \nabla PD^\epsilon(r_h(u^k)), \tilde{u}) \\ &\quad + \Delta t (\tau_h^k(v), \tilde{u}) + \Delta t (\sigma_{per,h}^k(u), \tilde{u}). \end{aligned} \tag{76}$$

Since u, v are C^2 in time, we can easily show

$$\|\tau_h^k(u)\| \leq \Delta t \sup_t \left\| \frac{\partial^2 u}{\partial t^2} \right\| \quad \text{and} \quad \|\tau_h^k(v)\| \leq \Delta t \sup_t \left\| \frac{\partial^2 v}{\partial t^2} \right\|.$$

To estimate $\sigma_{per,h}^k(u)$, we note the Lipschitz property of peridynamics force in L^2 norm (see Theorem 1). This leads us to

$$\|\sigma_{per,h}^k(u)\| \leq \frac{L}{\epsilon^2} \|u^k - r_h(u^k)\| \leq \frac{Lc}{\epsilon^2} h^2 \sup_t \|u(t)\|_2, \tag{77}$$

where we have relabeled the L^2 Lipschitz constant L_1 as L .

Let τ be given by

$$\begin{aligned} \tau &:= \sup_k \left(\|\tau_h^k(u)\| + \|\tau_h^k(v)\| + \|\sigma_{per,h}^k(u)\| \right) \\ &\leq C_t \Delta t + C_s \frac{h^2}{\epsilon^2}. \end{aligned} \tag{78}$$

where

$$C_t := \left\| \frac{\partial^2 u}{\partial t^2} \right\| + \left\| \frac{\partial^2 v}{\partial t^2} \right\| \quad \text{and} \quad C_s := Lc \sup_t \|u(t)\|_2. \tag{79}$$

In equation for $e_h^k(u)$ (see (74)), we take $\tilde{u} = e_h^{k+1}(u)$. Note that $e_h^{k+1}(u) = u_h^k - r_h(u^k) \in V_h$. We have

$$\|e_h^{k+1}(u)\|^2 = (e_h^k(u), e_h^{k+1}(u)) + \Delta t (e_h^{k+1}(v), e_h^{k+1}(u)) + \Delta t (\tau_h^k(u), e_h^{k+1}(u)).$$

Using the fact that $(u, v) \leq \|u\| \|v\|$, we get

$$\begin{aligned} \|e_h^{k+1}(u)\|^2 &\leq \|e_h^k(u)\| \|e_h^{k+1}(u)\| + \Delta t \|e_h^{k+1}(v)\| \|e_h^{k+1}(u)\| \\ &\quad + \Delta t \|\tau_h^k(u)\| \|e_h^{k+1}(u)\|. \end{aligned}$$

Canceling $\|e_h^{k+1}(u)\|$ from both sides gives

$$\|e_h^{k+1}(u)\| \leq \|e_h^k(u)\| + \Delta t \|e_h^{k+1}(v)\| + \Delta t \|\tau_h^k(u)\|. \tag{80}$$

Similarly, if we choose $\tilde{u} = e_h^{k+1}(v)$ in (76) and use the steps similar to above, we get

$$\begin{aligned} \|e_h^{k+1}(v)\| &\leq \|e_h^k(v)\| + \Delta t \|-\nabla PD^\epsilon(u_h^k) + \nabla PD^\epsilon(r_h(u^k))\| \\ &\quad + \Delta t \left(\|\tau_h^k(v)\| + \|\sigma_{per,h}^k(u)\| \right). \end{aligned} \tag{81}$$

Using the Lipschitz property of the peridynamics force in L^2 , we have

$$\|-\nabla PD^\epsilon(u_h^k) + \nabla PD^\epsilon(r_h(u^k))\| \leq \frac{L}{\epsilon^2} \|u_h^k - r_h(u^k)\| = \frac{L}{\epsilon^2} \|e_h^k(u)\|. \tag{82}$$

After adding (80) and (81) and substituting (82), we get

$$\begin{aligned} \|e_h^{k+1}(u)\| + \|e_h^{k+1}(v)\| &\leq \|e_h^k(u)\| + \|e_h^k(v)\| + \Delta t \|e_h^{k+1}(v)\| \\ &\quad + \frac{L}{\epsilon^2} \Delta t \|e_h^k(u)\| + \Delta t \tau \end{aligned}$$

where τ is defined in (78).

Let $e^k := \|e_h^k(u)\| + \|e_h^k(v)\|$. Assuming $L/\epsilon^2 \geq 1$, we get

$$\begin{aligned} e^{k+1} &\leq e^k + \Delta t e^{k+1} + \Delta t \frac{L}{\epsilon^2} e^k + \Delta t \tau \\ \Rightarrow e^{k+1} &\leq \frac{1 + \Delta t L/\epsilon^2}{1 - \Delta t} e^k + \frac{\Delta t}{1 - \Delta t} \tau. \end{aligned}$$

Substituting e^k recursively in above equation, we get

$$e^{k+1} \leq \left(\frac{1 + \Delta t L / \epsilon^2}{1 - \Delta t} \right)^{k+1} e^0 + \frac{\Delta t}{1 - \Delta t} \tau \sum_{j=0}^k \left(\frac{1 + \Delta t L / \epsilon^2}{1 - \Delta t} \right)^{k-j}.$$

Noting $1/(1 - \Delta t) = 1 + \Delta t + \Delta t^2 + O(\Delta t^3)$,

$$\frac{1 + \Delta t L / \epsilon^2}{1 - \Delta t} \leq 1 + (1 + L / \epsilon^2) \Delta t + (1 + L / \epsilon^2) \Delta t^2 + O(L / \epsilon^2) O(\Delta t^3),$$

and $(1 + a \Delta t)^k \leq \exp[ka \Delta t] \leq \exp[Ta]$ for $a > 0$, we get

$$\begin{aligned} \left(\frac{1 + \Delta t L / \epsilon^2}{1 - \Delta t} \right)^k &\leq \exp[k \Delta t (1 + L / \epsilon^2) + k \Delta t^2 (1 + L / \epsilon^2) + k O(L / \epsilon^2) O(\Delta t^3)] \\ &\leq \exp[T(1 + L / \epsilon^2) + T \Delta t (1 + L / \epsilon^2) + O(TL / \epsilon^2) O(\Delta t^2)] \\ &= \exp[T(1 + L / \epsilon^2)(1 + \Delta t + O(\Delta t^2))]. \end{aligned}$$

Substituting above estimates, we can easily show that

$$\begin{aligned} e^{k+1} &\leq \exp[T(1 + L / \epsilon^2)(1 + \Delta t + O(\Delta t^2))] \\ &\quad \left[e^0 + \Delta t (1 + \Delta t + O(\Delta t^2)) \tau \sum_{j=0}^k 1 \right] \\ &\leq \exp[T(1 + L / \epsilon^2)(1 + \Delta t + O(\Delta t^2))] [e^0 + k \Delta t (1 + \Delta t + O(\Delta t^2)) \tau]. \end{aligned}$$

Finally, we substitute above into (70) to have

$$\begin{aligned} E^k &\leq C_p h^2 + \exp[T(1 + L / \epsilon^2)(1 + \Delta t + O(\Delta t^2))] \\ &\quad [e^0 + k \Delta t (1 + \Delta t + O(\Delta t^2)) \tau]. \end{aligned}$$

After taking sup over $k \leq T / \Delta t$, we get the desired result and proof of Theorem 2 is complete.

We now consider the stability of linearized peridynamics model.

Stability Condition for Linearized Peridynamics

In this section, we linearize the peridynamics model and obtain a CFL-like stability condition. For problems where strains are small, the stability condition for the linearized model is expected to work for nonlinear model. The slope of peridynamics

potential f is constant for sufficiently small strain, and therefore for small strain, the nonlinear model behaves like a linear model. When displacement field is smooth, the difference between the linearized peridynamics force and nonlinear peridynamics force is of the order of ϵ . See [Proposition 4, Jha and Lipton 2017c].

In (5), linearization gives

$$-\nabla PD_l^\epsilon(u)(x) = \frac{4}{\epsilon^{d+1}\omega_d} \int_{H_\epsilon(x)} \omega(x)\omega(y)J^\epsilon(|y-x|)f'(0)S(u)e_{y-x}dy. \quad (83)$$

The corresponding bilinear form is denoted as a_l^ϵ and is given by

$$a_l^\epsilon(u, v) = \frac{2}{\epsilon^{d+1}\omega_d} \int_D \int_{H_\epsilon(x)} \omega(x)\omega(y)J^\epsilon(|y-x|)f'(0)|y-x|S(u)S(v)dydx. \quad (84)$$

We have

$$(-\nabla PD_l^\epsilon(u), v) = -a_l^\epsilon(u, v).$$

We now discuss the stability of the FEM approximation to the linearized problem. We replace $-\nabla PD^\epsilon$ by its linearization denoted by $-\nabla PD_l^\epsilon$ in (66) and (67). The corresponding approximate solution in V_h is denoted by $u_{l,h}^k$ where

$$\left(\frac{u_{l,h}^{k+1} - 2u_{l,h}^k + u_{l,h}^{k-1}}{\Delta t^2}, \tilde{u} \right) = (-\nabla PD_l^\epsilon(u_{l,h}^k), \tilde{u}) + (b_h^k, \tilde{u}), \quad \forall \tilde{u} \in V_h \quad (85)$$

and

$$\left(\frac{u_{l,h}^1 - u_{l,h}^0}{\Delta t^2}, \tilde{u} \right) = \frac{1}{2}(-\nabla PD^\epsilon(u_{l,h}^0), \tilde{u}) + \frac{1}{\Delta t}(v_{l,h}^0, \tilde{u}) + \frac{1}{2}(b_h^0, \tilde{u}), \quad \forall \tilde{u} \in V_h. \quad (86)$$

We will adopt the following notations

$$\begin{aligned} \bar{u}_h^{k+1} &:= \frac{u_h^{k+1} + u_h^k}{2}, \quad \bar{u}_h^k := \frac{u_h^k + u_h^{k-1}}{2}, \\ \bar{\partial}_t u_h^k &:= \frac{u_h^{k+1} - u_h^{k-1}}{2\Delta t}, \quad \bar{\partial}_t^+ u_h^k := \frac{u_h^{k+1} - u_h^k}{\Delta t}, \quad \bar{\partial}_t^- u_h^k := \frac{u_h^k - u_h^{k-1}}{\Delta t}. \end{aligned} \quad (87)$$

With above notations, we have

$$\bar{\partial}_t u_h^k = \frac{\bar{\partial}_t^+ u_h^k + \bar{\partial}_t^- u_h^k}{2} = \frac{\bar{u}_h^{k+1} - \bar{u}_h^k}{\Delta t}.$$

We also define

$$\bar{\partial}_{tt}u_h^k := \frac{u_h^{k+1} - 2u_h^k + u_h^{k-1}}{\Delta t^2} = \frac{\bar{\partial}_t^+ u_h^k - \bar{\partial}_t^- u_h^k}{\Delta t}.$$

We introduce the discrete energy associated with $u_{i,h}^k$ at time step k as defined by

$$\mathcal{E}(u_{i,h}^k) := \frac{1}{2} \left[\|\bar{\partial}_t^+ u_{i,h}^k\|^2 - \frac{\Delta t^2}{4} a_l^\epsilon(\bar{\partial}_t^+ u_{i,h}^k, \bar{\partial}_t^+ u_{i,h}^k) + a_l^\epsilon(\bar{u}_{i,h}^{k+1}, \bar{u}_{i,h}^{k+1}) \right]$$

Following [Theorem 4.1, Karaa 2012], we have

Theorem 6 (Stability of the central difference approximation of linearized peridynamics).

Let $u_{i,h}^k$ be the approximate solution of (85) and (86) with respect to linearized peridynamics. In the absence of the body force $b(t) = 0$ and for all t , if Δt satisfies the CFL-like condition

$$\frac{\Delta t^2}{4} \sup_{u \in V_h \setminus \{0\}} \frac{a_l^\epsilon(u, u)}{(u, u)} \leq 1, \tag{88}$$

then the discrete energy is positive and satisfies

$$\mathcal{E}(u_{i,h}^k) = \mathcal{E}(u_{i,h}^{k-1}), \tag{89}$$

and we have the stability

$$\mathcal{E}(u_{i,h}^k) = \mathcal{E}(u_{i,h}^0). \tag{90}$$

Proof. Set $b(t) = 0$. Noting that a_l^ϵ is bilinear, after adding and subtracting term $(\Delta t^2/4)a_l^\epsilon(\bar{\partial}_{tt}u_{i,h}^k, \tilde{u})$ to (85), and noting the following

$$u_{i,h}^k + \frac{\Delta t^2}{4} \bar{\partial}_{tt}u_{i,h}^k = \frac{\bar{u}_{i,h}^{k+1}}{2} + \frac{\bar{u}_{i,h}^k}{2}$$

we get

$$(\bar{\partial}_{tt}u_{i,h}^k, \tilde{u}) - \frac{\Delta t^2}{4} a_l^\epsilon(\bar{\partial}_{tt}u_{i,h}^k, \tilde{u}) + \frac{1}{2} a_l^\epsilon(\bar{u}_{i,h}^{k+1} + \bar{u}_{i,h}^k, \tilde{u}) = 0.$$

We let $\tilde{u} = \bar{\partial}_t u_{i,h}^k$, to write

$$(\bar{\partial}_{tt}u_{i,h}^k, \bar{\partial}_t u_{i,h}^k) - \frac{\Delta t^2}{4} a_l^\epsilon(\bar{\partial}_{tt}u_{i,h}^k, \bar{\partial}_t u_{i,h}^k) + \frac{1}{2} a_l^\epsilon(\bar{u}_{i,h}^{k+1} + \bar{u}_{i,h}^k, \bar{\partial}_t u_{i,h}^k) = 0.$$

It is easily shown that

$$\begin{aligned} (\bar{\partial}_t u_{l,h}^k, \bar{\partial}_t u_{l,h}^k) &= \left(\frac{\bar{\partial}_t^+ u_{l,h}^k - \bar{\partial}_t^- u_{l,h}^k}{\Delta t}, \frac{\bar{\partial}_t^+ u_{l,h}^k + \bar{\partial}_t^- u_{l,h}^k}{2} \right) \\ &= \frac{1}{2\Delta t} [\|\bar{\partial}_t^+ u_{l,h}^k\|^2 - \|\bar{\partial}_t^- u_{l,h}^k\|^2] \end{aligned}$$

and

$$a_l^\epsilon(\bar{\partial}_t u_{l,h}^k, \bar{\partial}_t u_{l,h}^k) = \frac{1}{2\Delta t} [a_l^\epsilon(\bar{\partial}_t^+ u_{l,h}^k, \bar{\partial}_t^+ u_{l,h}^k) - a_l^\epsilon(\bar{\partial}_t^- u_{l,h}^k, \bar{\partial}_t^- u_{l,h}^k)].$$

Noting that $\bar{\partial}_t u_{l,h}^k = (\bar{u}_{l,h}^{k+1} - \bar{u}_{l,h}^k)/\Delta t$, we get

$$\begin{aligned} &\frac{1}{2\Delta t} a_l^\epsilon(\bar{u}_{l,h}^{k+1} + \bar{u}_{l,h}^k, \bar{u}_{l,h}^{k+1} - \bar{u}_{l,h}^k) \\ &= \frac{1}{2\Delta t} [a_l^\epsilon(\bar{u}_{l,h}^{k+1}, \bar{u}_{l,h}^{k+1}) - a_l^\epsilon(\bar{u}_{l,h}^k, \bar{u}_{l,h}^k)]. \end{aligned}$$

After combining the above equations, we get

$$\begin{aligned} &\frac{1}{\Delta t} \left[\left(\frac{1}{2} \|\bar{\partial}_t^+ u_{l,h}^k\|^2 - \frac{\Delta t^2}{8} a_l^\epsilon(\bar{\partial}_t^+ u_{l,h}^k, \bar{\partial}_t^+ u_{l,h}^k) + \frac{1}{2} a_l^\epsilon(\bar{u}_{l,h}^{k+1}, \bar{u}_{l,h}^{k+1}) \right) \right. \\ &\quad \left. - \left(\frac{1}{2} \|\bar{\partial}_t^- u_{l,h}^k\|^2 - \frac{\Delta t^2}{8} a_l^\epsilon(\bar{\partial}_t^- u_{l,h}^k, \bar{\partial}_t^- u_{l,h}^k) + \frac{1}{2} a_l^\epsilon(\bar{u}_{l,h}^k, \bar{u}_{l,h}^k) \right) \right] = 0. \quad (91) \end{aligned}$$

We recognize the first term in bracket as $\mathcal{E}(u_{l,h}^k)$. We next prove that the second term is $\mathcal{E}(u_{l,h}^{k-1})$. We substitute $k = k - 1$ in the definition of $\mathcal{E}(u_{l,h}^k)$ to get

$$\mathcal{E}(u_{l,h}^{k-1}) = \frac{1}{2} \left[\|\bar{\partial}_t^+ u_{l,h}^{k-1}\|^2 - \frac{\Delta t^2}{4} a_l^\epsilon(\bar{\partial}_t^+ u_{l,h}^{k-1}, \bar{\partial}_t^+ u_{l,h}^{k-1}) + a_l^\epsilon(\bar{u}_{l,h}^k, \bar{u}_{l,h}^k) \right].$$

We clearly have $\bar{\partial}_t^+ u_{l,h}^{k-1} = \frac{u_{l,h}^{k-1+1} - u_{l,h}^{k-1}}{\Delta t} = \bar{\partial}_t^- u_{l,h}^k$, and this implies that the second term in (91) is $\mathcal{E}(u_{l,h}^{k-1})$. It now follows from (91) that $\mathcal{E}(u_{l,h}^k) = \mathcal{E}(u_{l,h}^{k-1})$.

The stability condition is such that discrete energy is positive. In the definition of $\mathcal{E}(u_{l,h}^k)$, we see that the second term is negative. We now obtain a condition on the time step that insures the sum of the first two terms is positive, and this will establish the positivity of $\mathcal{E}(u_{l,h}^k)$. Let $v = \bar{\partial}_t^+ u_{l,h}^k \in V_h$, and then we require

$$\|v\|^2 - \frac{\Delta t^2}{4} a_l^\epsilon(v, v) \geq 0 \quad \Rightarrow \quad \frac{\Delta t^2}{4} \frac{a_l^\epsilon(v, v)}{\|v\|^2} \leq 1 \quad (92)$$

Clearly if Δt satisfies

$$\frac{\Delta t^2}{4} \sup_{v \in V_h \setminus \{0\}} \frac{a_l^\epsilon(v, v)}{\|v\|^2} \leq 1 \quad (93)$$

then (92) is also satisfied and the discrete energy is positive. Iteration gives $\mathcal{E}(u_{l,h}^k) = \mathcal{E}(u_{l,h}^0)$ and the theorem is proved.

Conclusion

In this chapter we computed the a priori error incurred in finite element and finite difference discretizations of peridynamics. We show that for finite element approximation with linear elements, the rate of convergence is better as compared to rate of convergence of finite difference approximation. A CFL-like condition for the stability of linearized peridynamics is obtained. For the fully nonlinear problem, we find that for the semi-discrete approximation the energy at any instant is bounded by initial energy and work done by the body force.

This model has been analyzed using a quadrature-based finite element approximation in detail in Jha and Lipton (2017c) for nonlinear nonlocal models and their linearization assuming an a priori higher regularity of solutions. If one assumes more regular solutions with three continuous spatial derivatives (no cracks), then solutions of the nonlinear nonlocal model converge to those of the classical local elastodynamic model at the rate ϵ uniformly in time in the H^1 norm (see (Theorem 5, Jha and Lipton 2017c)). The numerical simulation of problems using finite differences for this model is carried out in Lipton et al. (2016) and Diehl et al. (2016). In earlier work (Tian and Du 2014) develop a framework for asymptotically compatible finite element schemes for linear problems where the solutions of the nonlocal problem are known to converge to a unique solution of the local problem. For the problems treated there, the discrete approximations associated with asymptotically compatible schemes converge if $h \rightarrow 0$ and $\epsilon \rightarrow 0$.

For the bond-based prototypical microelastic brittle material model analyzed here, the uniqueness property for the $\epsilon = 0$ problem is much less clear. The nonlinear nonlocal model treated in this chapter is an evolution in taking values in the vector space L^2 and can be identified with a sharp fracture evolution as $\epsilon \rightarrow 0$ (see Lipton 2014, 2016). The limit evolution is shown to be an element of the vector space, the space of special functions of bounded deformation referred to as SBD. The description and properties of this vector space can be found in Ambrosio et al. (1997). Unlike the linear nonlocal models, we do not necessarily have a unique sharp fracture evolution in the $\epsilon = 0$ limit. The uniqueness of the limit evolution for the nonlocal nonlinear model is an open question and remains to be established. The issue of nonuniqueness arises as the limit evolution is not completely characterized. What is currently missing is a limiting kinetic relation relating crack growth to crack driving force. Future work will seek to account for the missing information and address the issue of uniqueness for the limit problem.

Acknowledgements This material is based upon work supported by the US Army Research Laboratory and the US Army Research Office under contract/grant number W911NF1610456.

References

- A. Agwai, I. Guven, E. Madenci, Predicting crack propagation with peridynamics: a comparative study. *Int. J. Fract.* **171**(1), 65–78 (2011)
- B. Aksoylu, T. Mengesha, Results on nonlocal boundary value problems. *Numer. Funct. Anal. Optim.* **31**(12), 1301–1317 (2010)
- B. Aksoylu, M.L. Parks, Variational theory and domain decomposition for nonlocal problems. *Appl. Math. Comput.* **217**(14), 6498–6515 (2011)
- B. Aksoylu, Z. Unlu, Conditioning analysis of nonlocal integral operators in fractional sobolev spaces. *SIAM J. Numer. Anal.* **52**, 653–677 (2014)
- L. Ambrosio, A. Coscia, G. Dal Maso, Fine properties of functions with bounded deformation. *Arch. Ration. Mech. Anal.* **139**, 201–238 (1997)
- D.N. Arnold, Lecture notes on numerical analysis of partial differential equations (2011), <http://www.math.umn.edu/~arnold/8445/notes.pdf>
- E. Askari, F. Bobaru, R. Lehoucq, M. Parks, S. Silling, O. Weckner, Peridynamics for multiscale materials modeling. *J Phys Conf Ser* **125**, 012078 (2008). IOP Publishing
- G.A. Baker, Error estimates for finite element methods for second order hyperbolic equations. *SIAM J. Numer. Anal.* **13**(4), 564–576 (1976)
- F. Bobaru, W. Hu, The meaning, selection, and use of the peridynamic horizon and its relation to crack branching in brittle materials. *Int. J. Fract.* **176**(2), 215–222 (2012)
- F. Bobaru, M. Yang, L.F. Alves, S.A. Silling, E. Askari, J. Xu, Convergence, adaptive refinement, and scaling in 1d peridynamics. *Int. J. Numer. Meth. Eng.* **77**(6), 852–877 (2009)
- X. Chen, M. Gunzburger, Continuous and discontinuous finite element methods for a peridynamics model of mechanics. *Comput. Meth. Appl. Mech. Eng.* **200**(9), 1237–1250 (2011)
- K. Dayal, Leading-order nonlocal kinetic energy in peridynamics for consistent energetics and wave dispersion. *J. Mech. Phys. Solids* **105**, 235–253 (2017)
- P. Diehl, R. Lipton, M. Schweitzer, Numerical verification of a bond-based softening peridynamic model for small displacements: deducing material parameters from classical linear theory. Institut für Numerische Simulation Preprint, (2016)
- Q. Du, M. Gunzburger, R. Lehoucq, K. Zhou, Analysis of the volume-constrained peridynamic navier equation of linear elasticity. *J. Elast.* **113**(2), 193–217 (2013a)
- Q. Du, L. Tian, X. Zhao, A convergent adaptive finite element algorithm for nonlocal diffusion and peridynamic models. *SIAM J. Numer. Anal.* **51**(2), 1211–1234 (2013b)
- Q. Du, K. Zhou, Mathematical analysis for the peridynamic nonlocal continuum theory. *ESAIM Math. Model. Numer. Anal.* **45**(2), 217–234 (2011)
- Q. Du, Y. Tao, X. Tian, A peridynamic model of fracture mechanics with bond-breaking. *J. Elast.* (2017). <https://doi.org/10.1007/s10659-017-9661-2>
- E. Emmrich, R.B. Lehoucq, D. Puhst, Peridynamics: a nonlocal continuum theory, in *Meshfree Methods for Partial Differential Equations VI* (Springer, Berlin/Heidelberg, 2013), pp. 45–65
- E. Emmrich, O. Weckner, et al. On the well-posedness of the linear peridynamic model and its convergence towards the navier equation of linear elasticity. *Commun. Math. Sci.* **5**(4), 851–864 (2007)
- E. Emmrich, D. Puhst, A short note on modeling damage in peridynamics. *J. Elast.* **123**, 245–252 (2016)
- J.T. Foster, S.A. Silling, W. Chen, An energy based failure criterion for use with peridynamic states. *Int. J. Multiscale Comput. Eng.* **9**(6), 675–688 (2011)
- W. Gerstle, N. Sau, S. Silling, Peridynamic modeling of concrete structures. *Nucl. Eng. Des.* **237**(12), 1250–1258 (2007)

- M.J. Grote, D. Schötzau, Optimal error estimates for the fully discrete interior penalty dg method for the wave equation. *J. Sci. Comput.* **40**(1), 257–272 (2009)
- Q. Guan, M. Gunzburger, Stability and accuracy of time-stepping schemes and dispersion relations for a nonlocal wave equation. *Numer. Meth. Partial Differ. Equ.* **31**(2), 500–516 (2015)
- Y.D. Ha, F. Bobaru, Characteristics of dynamic brittle fracture captured with peridynamics. *Eng. Fract. Mech.* **78**(6), 1156–1168 (2011)
- P.K. Jha, R. Lipton, Finite element approximation of nonlocal fracture models (2017a). arXiv preprint arXiv:1710.07661
- P.K. Jha, R. Lipton, Numerical analysis of nonlocal fracture models models in holder space (2017b). arXiv preprint arXiv:1701.02818. To appear in *SIAM Journal on Numerical Analysis* 2018
- P.K. Jha, R. Lipton, Numerical convergence of nonlinear nonlocal continuum models to local elastodynamics (2017c). arXiv preprint arXiv:1707.00398. To appear in *International Journal for Numerical Methods in Engineering* 2018
- S. Karaa, Stability and convergence of fully discrete finite element schemes for the acoustic wave equation. *J. Appl. Math. Comput.* **40**(1–2), 659–682 (2012)
- R. Lipton, Dynamic brittle fracture as a small horizon limit of peridynamics. *J. Elast.* **117**(1), 21–50 (2014)
- R. Lipton, Cohesive dynamics and brittle fracture. *J. Elast.* **124**(2), 143–191 (2016)
- R. Lipton, S. Silling, R. Lehoucq, Complex fracture nucleation and evolution with nonlocal elastodynamics (2016). arXiv preprint arXiv:1602.00247
- R. Lipton, E. Said, P. Jha, Free damage propagation with memory. *Journal of Elasticity*, 1–25 (2018). <https://doi.org/10.1007/s10659-018-9672>
- D.J. Littlewood, Simulation of dynamic fracture using peridynamics, finite element modeling, and contact, in *Proceedings of the ASME 2010 International Mechanical Engineering Congress and Exposition (IMECE)* (2010)
- R.W. Macek, S.A. Silling, Peridynamics via finite element analysis. *Finite Elem. Anal. Des.* **43**(15), 1169–1178 (2007)
- T. Mengesha, Q. Du, Analysis of a scalar peridynamic model with a sign changing kernel. *Discrete Contin. Dynam. Syst. B* **18**, 1415–1437 (2013)
- T. Mengesha, Q. Du, Nonlocal constrained value problems for a linear peridynamic navier equation. *J. Elast.* **116**(1), 27–51 (2014)
- S. Silling, O. Weckner, E. Askari, F. Bobaru, Crack nucleation in a peridynamic solid. *Int. J. Fract.* **162**(1–2), 219–227 (2010)
- S.A. Silling, Reformulation of elasticity theory for discontinuities and long-range forces. *J. Mech. Phys. Solids* **48**(1), 175–209 (2000)
- X. Tian, Q. Du, Asymptotically compatible schemes and applications to robust discretization of nonlocal models. *SIAM J. Numer. Anal.* **52**(4), 1641–1665 (2014)
- X. Tian, Q. Du, M. Gunzburger, Asymptotically compatible schemes for the approximation of fractional laplacian and related nonlocal diffusion problems on bounded domains. *Adv. Comput. Math.* **42**(6), 1363–1380 (2016a)
- X. Tian, Q. Du, M. Gunzburger, Asymptotically compatible schemes for the approximation of fractional laplacian and related nonlocal diffusion problems on bounded domains. *Adv. Comput. Math.* **42**(6), 1363–1380 (2016b)
- O. Weckner, E. Emmrich, Numerical simulation of the dynamics of a nonlocal, inhomogeneous, infinite bar. *J. Comput. Appl. Mech.* **6**(2), 311–319 (2005)
- G. Zhang, Q. Le, A. Loghin, A. Subramaniyan, F. Bobaru, Validation of a peridynamic model for fatigue cracking. *Eng. Fract. Mech.* **162**, 76–94 (2016)



Dynamic Damage Propagation with Memory: A State-Based Model

45

Robert Lipton, Eyad Said, and Prashant K. Jha

Contents

Introduction	1496
Formulation	1497
Existence of Solutions	1501
Energy Balance	1508
Explicit Damage Models, Cyclic Loading, and Strain to Failure	1512
Numerical Results	1516
Periodic Loading	1516
Shear Loading	1517
Linear Elastic Operators in the Small Horizon Limit	1519
Conclusions	1521
References	1522

Abstract

A model for dynamic damage propagation is developed using nonlocal potentials. The model is posed using a state-based peridynamic formulation. The resulting evolution is seen to be well posed. At each instant of the evolution, we identify a damage set. On this set, the local strain has exceeded critical values either for tensile or hydrostatic strain, and damage has occurred. The damage set is nondecreasing with time and is associated with damage state variables defined at each point in the body. We show that a rate form of energy balance holds at

R. Lipton (✉)

Department of Mathematics and Center for Computation and Technology, Louisiana State University, Baton Rouge, LA, USA
e-mail: lipton@lsu.edu

E. Said · P. K. Jha

Department of Mathematics, Louisiana State University, Baton Rouge, LA, USA
e-mail: esaid1@lsu.edu; prashant.j16o@gmail.com

each time during the evolution. Away from the damage set, we show that the nonlocal model converges to the linear elastic model in the limit of vanishing nonlocal interaction.

Keywords

Damage model · Nonlocal interactions · Energy dissipation · State-based peridynamics

Introduction

In this chapter, we address the problem of damage propagation in materials. Here the damage evolution is not known a priori and is found as part of the problem solution. Our approach is to use a nonlocal formulation with the purpose of using the least number of parameters to describe the model. We will work within the small deformation setting, and the model is developed within a state-based peridynamic formulation. Here strains are expressed in terms of displacement differences as opposed to spatial derivatives. For the problem at hand, the nonlocality provides the flexibility to simultaneously model non-differentiable displacements and damage evolution. The net force acting on a point x is due to the strain between x and neighboring points y . The neighborhood of nonlocal interaction between x and its neighbors y is confined to ball of radius δ centered at x denoted by $B_\delta(x)$. The radius of the ball is called the horizon. Numerical implementations based on nonlocal peridynamic models exhibit formation and localization of features associated with phase transformation and fracture (see, e.g., Dayal and Bhattacharya 2006; Silling and Lehoucq 2008; Silling et al. 2010; Foster et al. 2011; Agwai et al. 2011; Lipton et al. 2016; Bobaru and Hu 2012; Ha and Bobaru 2010; Silling and Bobaru 2005; Weckner and Abeyaratne 2005; Gerstle et al. 2007; Weckner and Emmrich 2005). A recent review can be found in Bobaru et al. (2016).

The recent model studied in Lipton (2014, 2016), Lipton et al. (2016), and Jha and Lipton (2017) is defined by double-well two-point strain potentials. Here one potential well is centered at the origin and associated with elastic response, while the other well is at infinity and associated with surface energy. The rationale for studying these models is that they are shown to be well posed, and in the limit of vanishing nonlocality, the dynamics recovers features associated with sharp fracture propagation (see Lipton 2014, 2016). While memory is not incorporated in this model, it is seen that the inertia of the evolution keeps the forces in a softened state over time as evidenced in simulations (Lipton et al. 2016). This modeling approach is promising for fast cracks, but for cyclic loading and slowly propagating fractures, an explicit damage-fracture modeling with memory is needed. In this work, we develop this approach for more general models that allow for three-point nonlocal interactions and irreversible damage. The use of three-point potentials allows one to model a larger variety of elastic properties. In the lexicon of peridynamics, we adopt an ordinary state-based formulation (Silling 2000; Silling et al. 2007). We introduce nonlocal forces that soften irreversibly as the shear strain or dilatational strain

increases beyond critical values. This model is shown to deliver a mathematically well-posed evolution. Our proof of this is motivated by recent work Emrich and Phulst (2016) where existence of solution for bond-based peridynamic models with damage is established. Recently another well-posed bond-based model with damage has been proposed in Du et al. (2017) where fracture simulations are carried out.

In addition to being state based, our modeling approach differs from Emrich and Phulst (2016) and earlier bond-based work Silling and Askari (2005) and uses differentiable damage variables. This feature allows us to establish an energy balance equation relating kinetic energy, potential energy, and energy dissipation at each instant during the evolution. At each instant, we identify the set undergoing damage where the local energy dissipation rate is positive. On this set, the local strain has exceeded a critical value, and damage has occurred. Damage is irreversible, and the damage set is monotonically increasing with time. Explicit damage models are illustrated, and stress strain curves for both cyclic loading and strain to failure are provided. These models are illustrated in two numerical examples. In the first example, we consider a square domain and apply a time periodic y -directed displacement along the top edge while fixing the bottom, left and right edges. We track the strain and force over three loading periods. The simulations show that bonds suffer damage and the strain vs force plot is similar to the one predicted by the damage law (see Fig. 14). In the second example, we apply a shear load to the top edge while fixing the bottom edge and leaving left and right edges free. As expected, we find that damage appears along the diagonal of square (see Fig. 15).

We conclude by noting that for this model the forces scale inversely with the length of the horizon. With this in mind, we consider undamaged regions, and we are able to show that the nonlocal operator converges to a linear local operator associated with the elastic wave equation. In this limit, the elastic tensor can have any combination of Poisson's ratio and Young's modulus. The Poisson's ratio and Young modulus are determined uniquely by explicit formulas in terms of the nonlocal potentials used to define the model. This result is consistent with small horizon convergence results for convex energies (see Emrich and Weckner 2007; Mengesha and Du 2014; Silling and Lehoucq 2008). Further reading and complete derivations can be found in the recent monograph Lipton et al. (2018).

Formulation

In this work, we assume the displacements u are small (infinitesimal) relative to the size of the three-dimensional body D . The tensile strain is denoted $S = S(y, x, t; u)$ and given by

$$S(y, x, t; u) = \frac{u(t, y) - u(t, x)}{|y - x|} \cdot e_{y-x}, \quad e_{y-x} = \frac{y - x}{|y - x|}, \quad (1)$$

where e_{y-x} is a unit direction vector and \cdot is the dot product. It is evident that $S(y, x, t; u)$ is the tensile strain along the direction e_{y-x} . We introduce the nonnegative weight $\omega^\delta(|y-x|)$ such that $\omega^\delta = 0$ for $|y-x| > \delta$ and the hydrostatic strain at x is defined by

$$\theta(x, t; u) = \frac{1}{V_\delta} \int_{D \cap B_\delta(x)} \omega^\delta(|y-x|) |y-x| S(y, x, t; u) dy, \tag{2}$$

where V_δ is the volume of the ball $B_\delta(x)$ of radius δ centered at x . The weight is chosen such that $\omega^\delta(|y-x|) = \omega(|y-x|/\delta)$ and

$$\ell_1 = \frac{1}{V_\delta} \int_{B_\delta(x)} \omega^\delta(|y-x|) dy < \infty. \tag{3}$$

We follow Silling (2000) and Emrich and Phulst (2016) and introduce a nonnegative damage factor taking the value one in the undamaged region and zero in the fully damaged region. The damage factor for the force associated with tensile strains is written $H^T(u)(y, x, t)$; the corresponding factor for hydrostatic strains is written $H^D(u)(x, t)$. Here we assume no damage and $H^T(u)(y, x, t) = 1$ until a critical tensile strain S_c is reached. For tensile strains greater than S_c , damage is initiated and $H^T(y, x, t)$ drops below 1. The fully damaged state is $H^T(y, x, t) = 0$. For hydrostatic strains, we assume no damage until a critical positive dilatational strain θ_c^+ or a negative compressive strain (θ_c^-) is reached. Again $H^D(x, t) = 1$ until a critical hydrostatic strain is reached and then drops below 1 with the fully damaged state being $H^D(x, t) = 0$. We postpone description of the specific form of the history-dependent damage factors until after we have defined the nonlocal forces.

The force at a point x due to tensile strain is given by

$$\begin{aligned} \mathcal{L}^T(u)(x, t) &= \frac{2}{V_\delta} \int_{D \cap B_\delta(x)} \frac{J^\delta(|y-x|)}{\delta|y-x|} H^T(u)(y, x, t) \partial_S \\ & f(\sqrt{|y-x|} S(y, x, t; u)) e_{y-x} dy, \end{aligned} \tag{4}$$

Here $J^\delta(|y-x|)$ is a nonnegative bounded function such that $J^\delta = 0$ for $|y-x| > \delta$ and $M = \sup\{y \in B_\delta(x); J^\delta(|y-x|)\}$ and

$$\ell_2 = \frac{1}{V_\delta} \int_{B_\delta(x)} \frac{J^\delta(|y-x|)}{|y-x|^2} dy < \infty \text{ and } \ell_3 = \frac{1}{V_\delta} \int_{B_\delta(x)} \frac{J^\delta(|y-x|)}{|y-x|^{3/2}} dy < \infty. \tag{5}$$

Both J^δ and ω^δ are prescribed and characterize the influence of nonlocal forces on x by neighboring points y . Here ∂_S is the partial derivative with respect to strain. The function $f = f(r)$ is twice differentiable for all arguments r on the real line, and f' and f'' are bounded. Here we take $f(r) = \alpha r^2/2$ for $r < r_1$ and $f = r$ for

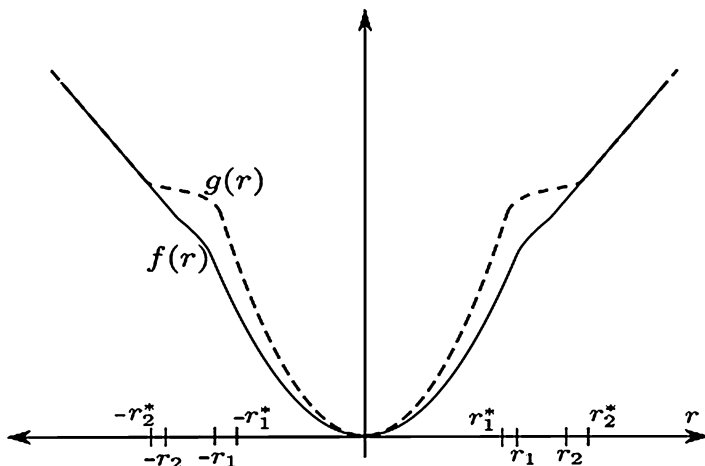


Fig. 1 Generic plot of $f(r)$ (Solid line) and $g(r)$ (Dashed line)

$r_2 < r$, with $r_1 < r_2$ (see Fig. 1). The factor $\sqrt{|y - x|}$ appearing in the argument of $\partial_S f$ ensures that the nonlocal operator \mathcal{L}^T converges to the divergence of a stress tensor in the small horizon limit when it's known a priori that displacements are smooth (see section “Linear Elastic Operators in the Small Horizon Limit”).

The force at a point x due to the hydrostatic strain is given by

$$\mathcal{L}^D(u)(x, t) = \frac{1}{V_\delta} \int_{D \cap B_\delta(x)} \frac{\omega^\delta(|y - x|)}{\delta^2} [H^D(u)(y, t) \partial_\theta g(\theta(y, t; u)) + \quad (6)$$

$$H^D(u)(x, t) \partial_\theta g(\theta(x, t; u))] e_{y-x} dy, \quad (7)$$

where the function $g(r) = \beta r^2/2$ for $r < r_1^*$, $g = r$ for $r_2^* < r$, with $r_1^* < r_2^*$ and g is twice differentiable and g' and g'' are bounded (see Fig. 1). It is readily verified that the force $\mathcal{L}^T(u)(x, t) + \mathcal{L}^D(u)(x, t)$ satisfies balance of linear and angular momentum.

The damage factor for tensile strain $H^T(u)(y, x, t)$ is given in terms of the functions $h(x)$ and $j_S(x)$. Here h is nonnegative, has bounded derivatives (hence Lipschitz continuous), takes the value one for negative x and for $x \geq 0$ decreases, and is zero for $x > x_c$ (see Fig. 2). Here we are free to choose x_c to be any small and positive number. The function $j_S(x)$ is nonnegative, has bounded derivatives (hence Lipschitz continuous), takes the value zero up to a positive critical strain S_C , and then takes on positive values. We will suppose $j_S(x) \leq \gamma|x|$ for some $\gamma > 0$ (see Fig. 3). The damage factor is now defined to be

$$H^T(u)(y, x, t) = h \left(\int_0^t j_S(S(y, x, \tau; u)) d\tau \right). \quad (8)$$

Fig. 2 Generic plot of $h(x)$

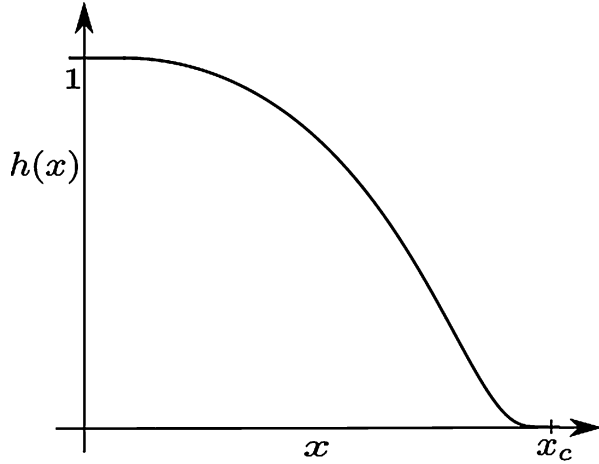
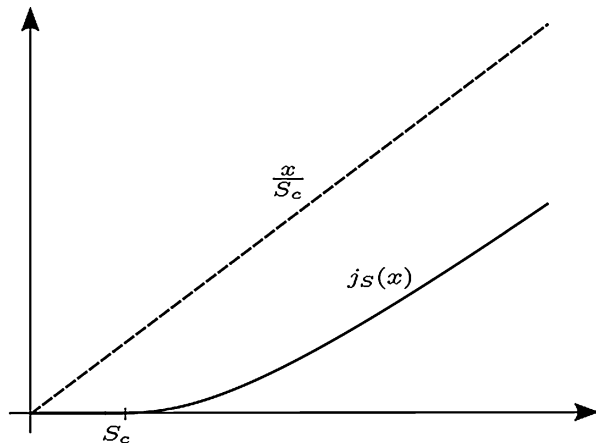


Fig. 3 Generic plot of $j_S(x)$ with S_c

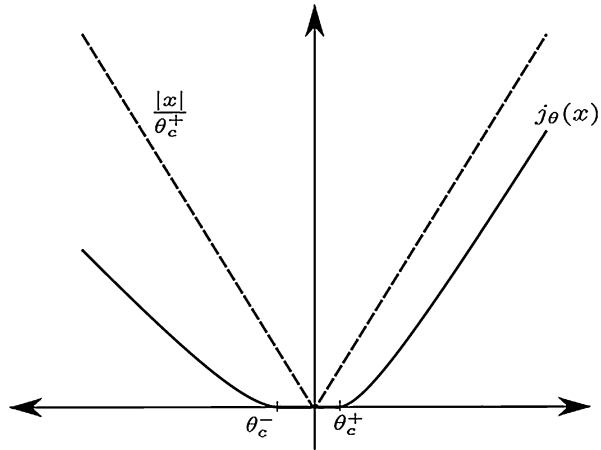


It is clear from this definition that damage occurs when the stress exceeds S_c for some period of time and the bond force decreases irrevocably from its undamaged value. The damage function defined here is symmetric, i.e., $H^T(u)(y, x, t) = H^T(u)(x, y, t)$. For hydrostatic strain, we introduce the nonnegative function j_θ with bounded derivatives (hence Lipschitz continuous). We suppose $j_\theta = 0$ for an interval containing the origin given by (θ_c^-, θ_c^+) and take positive values outside this interval (see Fig. 4). As before we will suppose $j_\theta(x) \leq \gamma|x|$ for some $\gamma > 0$. The damage factor for hydrostatic strain is given by

$$H^D(u)(x, t) = h \left(\int_0^t j_\theta(\theta(x, \tau; u)) d\tau \right). \tag{9}$$

For this model, it is clear that damage can occur irreversibly for compressive or dilatational strain when the possibly different critical values θ_c^- or θ_c^+ are exceeded.

Fig. 4 Generic plot of $j_\theta(x)$ with θ_c^+ , and θ_c^-



The *damage set* at time t is defined to be the collection of all points x for which $H^T(y, x, t)$ or $H^D(u)(x, t)$ is less than one. This set is monotonically increasing in time. The *process zone* at time t is the collection of points x undergoing damage such that $\partial_t H^T(y, x, t) < 0$ or $\partial_t H^D(x, t) < 0$. Explicit examples of $H^T(u)(y, x, t)$ and $H^D(u)(x, t)$ are given in section “[Explicit Damage Models, Cyclic Loading, and Strain to Failure.](#)”

We define the body force $b(x, t)$, and the displacement $u(x, t)$ is the solution of the initial value problem given by

$$\rho \partial_t^2 u(x, t) = \mathcal{L}^T(u)(x, t) + \mathcal{L}^D(u)(x, t) + b(x, t) \text{ for } x \in D \text{ and } t \in (0, T), \tag{10}$$

with initial data

$$u(x, 0) = u_0(x), \quad \partial_t u(x, 0) = v_0(x). \tag{11}$$

It is easily verified that this is an ordinary state-based peridynamic model. We show in the next section that this initial value problem is well posed.

Existence of Solutions

The regularity and existence of the solution depend on the regularity of the initial data and body force. In this work, we choose a general class of body forces and initial conditions. The initial displacement u_0 and velocity v_0 are chosen to be integrable and bounded and belonging to $L^\infty(D; \mathbb{R}^3)$. The space of such functions is denoted by $L^\infty(D; \mathbb{R}^3)$. The body force $b(x, t)$ is chosen such that for every $t \in [0, T_0]$, b takes values in $L^\infty(D, \mathbb{R}^3)$ and is continuous in time. The associated norm is defined to be $\|b\|_{C([0, T_0]; L^\infty(D, \mathbb{R}^3))} = \max_{t \in [0, T_0]} \|b(x, t)\|_{L^\infty(D, \mathbb{R}^3)}$. The

associated space of continuous functions in time taking values in $L^\infty(D; \mathbb{R}^3)$ for which this norm is finite is denoted by $C([0, T_0]; L^\infty(D, \mathbb{R}^3))$. The space of functions twice differentiable in time taking values in $L^\infty(D, \mathbb{R}^3)$ such that both derivatives belong to $C([0, T_0]; L^\infty(D, \mathbb{R}^3))$ is written as $C^2([0, T_0]; L^\infty(D, \mathbb{R}^3))$. We now assert the existence and uniqueness for the solution of the initial value problem.

Theorem 1 (Existence and uniqueness of the damage evolution). *The initial value problem given by (10) and (11) has a solution $u(x, t)$ such that for every $t \in [0, T_0]$, u takes values in $L^\infty(D, \mathbb{R}^3)$ and is the unique solution belonging to the space $C^2([0, T_0]; L^\infty(D, \mathbb{R}^3))$.*

To prove the theorem, we will show

- (1) The operator $\mathcal{L}^T(u)(x, t) + \mathcal{L}^D(u)(x, t)$ is a map from $C([0, T_0]; L^\infty(D, \mathbb{R}^3))$ into itself.
- (2) The operator $\mathcal{L}^T(u)(x, t) + \mathcal{L}^D(u)(x, t)$ is Lipschitz continuous with respect to the norm of $C([0, T_0]; L^\infty(D, \mathbb{R}^3))$.

The theorem then follows from an application of the Banach fixed point theorem.

To establish properties (1) and (2), we state and prove the following lemmas for the damage factors.

Lemma 1. *Let $H^T(u)(y, x, t)$ and $H^D(u)(x, t)$ be defined as in (8) and (9). Then for $u \in C([0, T_0]; L^\infty(D, \mathbb{R}^3))$, the mappings*

$$(y, x) \mapsto H^T(u)(y, x, t) : D \times D \rightarrow \mathbb{R}, \quad x \mapsto H^D(u)(x, t) : D \rightarrow \mathbb{R} \quad (12)$$

are measurable for every $t \in [0, T_0]$, and the mappings

$$t \mapsto H^T(u)(y, x, t) : [0, T_0] \rightarrow \mathbb{R}, \quad t \mapsto H^D(u)(x, t) : [0, T_0] \rightarrow \mathbb{R} \quad (13)$$

are continuous for almost all (y, x) and x , respectively. Moreover for almost all $(y, x) \in D \times D$ and all $t \in [0, T_0]$, the map

$$u \mapsto H^T(u)(y, x, t) : C([0, T_0]; L^\infty(D, \mathbb{R}^3)) \rightarrow \mathbb{R} \quad (14)$$

is Lipschitz continuous, and for almost all $x \in D$ and all $t \in [0, T_0]$, the map

$$u \mapsto H^D(u)(x, t) : C([0, T_0]; L^\infty(D, \mathbb{R}^3)) \rightarrow \mathbb{R} \quad (15)$$

is Lipschitz continuous.

Proof. The measurability properties are immediate. In what follows, constants are generic and apply to the context in which they are used. We establish continuity in time for $H^D(u)$. For \hat{t} and t in $[0, T_0]$, we have

$$\begin{aligned}
 & |H^D(u)(x, \hat{t}) - H^D(u)(x, t)| \\
 &= \left| h \left(\int_0^{\hat{t}} j_\theta(\theta(x, \tau; u)) d\tau \right) - h \left(\int_0^t j_\theta(\theta(x, \tau; u)) d\tau \right) \right| \\
 &\leq C_1 \int_{\min\{\hat{t}, t\}}^{\max\{\hat{t}, t\}} j_\theta(\theta(x, \tau; u)) d\tau \tag{16} \\
 &\leq \gamma C_1 \int_{\min\{\hat{t}, t\}}^{\max\{\hat{t}, t\}} |\theta(x, \tau; u)| d\tau \\
 &\leq \gamma \ell_1 C_1 C_2 |\hat{t} - t| 2 \|u\|_{C([0, T_0]; L^\infty(D, \mathbb{R}^3))}.
 \end{aligned}$$

The first inequality follows from the Lipschitz continuity of h , the second follows from the growth condition on j_θ , and the third follows from (3).

We establish continuity in time for $H^T(u)$. For \hat{t} and t in $[0, T_0]$, we have

$$\begin{aligned}
 & |H^T(u)(x, \hat{t}) - H^T(u)(x, t)| \\
 &= \left| h \left(\int_0^{\hat{t}} j_S(S(y, x, \tau; u)) d\tau \right) - h \left(\int_0^t j_S(S(y, x, \tau; u)) d\tau \right) \right| \\
 &\leq C_1 \int_{\min\{\hat{t}, t\}}^{\max\{\hat{t}, t\}} j_S(S(y, x, \tau; u)) d\tau \tag{17} \\
 &\leq \gamma C_1 \int_{\min\{\hat{t}, t\}}^{\max\{\hat{t}, t\}} |S(y, x, \tau; u)| d\tau \\
 &\leq \gamma C_1 C_2 \frac{|\hat{t} - t|}{|y - x|} 2 \|u\|_{C([0, T_0]; L^\infty(D, \mathbb{R}^3))}.
 \end{aligned}$$

The first inequality follows from the Lipschitz continuity of h , the second follows from the growth condition on j_S , and the third follows from the definition of strain (1).

To demonstrate Lipschitz continuity for $H^D(u)(x, t)$, we write

$$\begin{aligned}
 & |H^D(u)(x, t) - H^D(v)(x, t)| \\
 &= \left| h \left(\int_0^t j_\theta(\theta(x, \tau; u)) d\tau \right) - h \left(\int_0^t j_\theta(\theta(x, \tau; v)) d\tau \right) \right| \\
 &\leq C_1 \left| \int_0^t (j_\theta(\theta(x, \tau; u)) - j_\theta(\theta(x, \tau; v))) d\tau \right| \tag{18} \\
 &\leq C_1 C_2 \int_0^t |\theta(x, \tau; u) - \theta(x, \tau; v)| d\tau \\
 &\leq 2t \ell_1 C_1 C_2 \|u - v\|_{C([0, t]; L^\infty(D, \mathbb{R}^3))}.
 \end{aligned}$$

The first inequality follows from the Lipschitz continuity of h , the second follows from the Lipschitz continuity of j_θ , and the third follows from (3). The Lipschitz continuity for $H^S(u)(y, x, t)$ follows from similar arguments using the Lipschitz continuity of h , j_S , and (1), and we get

$$\begin{aligned} & |H^T(u)(y, x, t) - H^T(v)(y, x, t)| \\ & \leq \frac{2tC_1C_2C_3}{|y-x|} \|u-v\|_{C([0,t];L^\infty(D,\mathbb{R}^3))}. \end{aligned} \tag{19}$$

□

Proof of Theorem 1. We establish (1) by first noting that

$$|\mathcal{L}^T(u)(x, t) + \mathcal{L}^D(u)(x, t)| \leq \frac{C}{\delta^2}, \tag{20}$$

where C is a constant. This estimate follows from the boundedness of f' , g' , $H^T(u)$, and $H^D(u)$ and the integrability of the ratios $J^\delta(|y-x|)/|y-x|^2$, $J^\delta(|y-x|)/|y-x|^{3/2}$, and $\omega^\delta(|y-x|)$. Thus $\|\mathcal{L}^T(u)(x, t) + \mathcal{L}^D(u)(x, t)\|_{L^\infty(D,\mathbb{R}^3)}$ is uniformly bounded for all $t \in [0, T_0]$.

To complete the demonstration of (1), we point out that the force functions $\partial_S f$ and $\partial_\theta g$ are Lipschitz continuous in their arguments. The key features are given in the following lemma.

Lemma 2. *Given two functions v and w in $L^\infty(D, \mathbb{R}^3)$, then*

$$\begin{aligned} & |\partial_S f(\sqrt{|y-x|}S(y, x; v)) - \partial_S f(\sqrt{|y-x|}S(y, x; w))| \\ & \leq \frac{2C}{\sqrt{|y-x|}} \|v-w\|_{L^\infty(D,\mathbb{R}^3)}. \end{aligned} \tag{21}$$

and

$$|\partial_\theta g(\theta(x; v)) - \partial_\theta g(\theta(x; w))| \leq 2\ell_1 C \|v-w\|_{L^\infty(D,\mathbb{R}^3)}. \tag{22}$$

Proof.

$$\begin{aligned} & |\partial_S f(\sqrt{|y-x|}S(y, x; v)) - \partial_S f(\sqrt{|y-x|}S(y, x; w))| \\ & \leq C \sqrt{|y-x|} |S(y, x; v) - S(y, x; w)| \leq \frac{2C}{\sqrt{|y-x|}} \|v-w\|_{L^\infty(D,\mathbb{R}^3)}, \end{aligned} \tag{23}$$

where the first inequality follows from the Lipschitz continuity of $\partial_S f$ and the second follows from the definition of S .

For $\partial_\theta g$, we have

$$|\partial_\theta g(\theta(x; v)) - \partial_\theta g(\theta(x; w))| \leq C|\theta(x; v) - \theta(x; w)| \leq 2\ell_1 C_1 \|v - w\|_{L^\infty(D, \mathbb{R}^3)}, \tag{24}$$

where the first inequality follows from the Lipschitz continuity of $\partial_\theta g$ and the second follows from the definitions of θ and S . \square

We have

$$\begin{aligned} & |\mathcal{L}^T(u)(x, \hat{t}) - \mathcal{L}^T(u)(x, t)| \\ & \leq \frac{2}{V_\delta} \int_{D \cap B_\delta(x)} \frac{J^\delta(|y-x|)}{\delta|y-x|} |\partial_S f(\sqrt{y-x} S(y, x, \hat{t}; u)) \\ & \quad - \partial_S f(\sqrt{y-x} S(y, x, t; u))| dy \\ & \quad + \frac{2}{V_\delta} \int_{D \cap B_\delta(x)} \frac{J^\delta(|y-x|)}{\delta|y-x|} |H^T(u)(y, x, \hat{t}) - H^T(u)(y, x, t)| dy. \end{aligned} \tag{25}$$

From the above, (19), and Lemma 2, we see that

$$\begin{aligned} & \|\mathcal{L}^T(u)(x, \hat{t}) - \mathcal{L}^T(u)(x, t)\|_{L^\infty(D, \mathbb{R}^3)} \\ & \leq \frac{\ell_3 C_3}{\delta} \|u(x, \hat{t}) - u(x, t)\|_{L^\infty(D, \mathbb{R}^3)} + \frac{\ell_2 \gamma C_1 C_2}{\delta} |\hat{t} - t| 2\|u\|_{C([0, T_0]; L^\infty(D, \mathbb{R}^3))} \end{aligned} \tag{26}$$

and we see \mathcal{L}^T is well defined and maps $C([0, T_0]; L^\infty(D, \mathbb{R}^3))$ into itself.

We show the continuity in time for $\mathcal{L}^D(u)(x, t)$. Now we have

$$\begin{aligned} & |\mathcal{L}^D(u)(x, \hat{t}) - \mathcal{L}^D(u)(x, t)| \\ & \leq \frac{1}{V_\delta} \int_{D \cap B_\delta(x)} \frac{\omega^\delta(|y-x|)}{\delta^2} |\partial_\theta g(\theta(y, \hat{t}; u)) - \partial_\theta g(\theta(y, t; u))| dy \\ & \quad + \frac{1}{V_\delta} \int_{D \cap B_\delta(x)} \frac{\omega^\delta(|y-x|)}{\delta^2} |H^D(u)(y, \hat{t}) - H^D(u)(y, t)| dy \\ & \quad + \frac{1}{V_\delta} \int_{D \cap B_\delta(x)} \frac{\omega^\delta(|y-x|)}{\delta^2} |\partial_\theta g(\theta(x, \hat{t}; u)) - \partial_\theta g(\theta(x, t; u))| dy \\ & \quad + \frac{1}{V_\delta} \int_{D \cap B_\delta(x)} \frac{\omega^\delta(|y-x|)}{\delta^2} |H^D(u)(x, \hat{t}) - H^D(u)(x, t)| dy \end{aligned} \tag{27}$$

and applying Lemma 2 and (18), (19), (20), (21), (22), (23), (24), (25), (26), and (27), we get the continuity

$$\begin{aligned} & |\mathcal{L}^D(u)(x, \hat{t}) - \mathcal{L}^D(u)(x, t)| \leq \frac{4\ell_1^2 C_1}{\delta^2} \|u(\hat{t}, x) - u(t, x)\|_{L^\infty(D, \mathbb{R}^3)} \\ & \quad + \frac{\gamma 4\ell_1^2 C_1 C_2}{\delta^2} |\hat{t} - t| \|u\|_{C([0, T_0]; L^\infty(D, \mathbb{R}^3))}. \end{aligned} \tag{28}$$

We conclude that \mathcal{L}^D is well defined and maps $C([0, T_0]; L^\infty(D, \mathbb{R}^3))$ into itself and item (1) is proved.

To show Lipschitz continuity, consider any two functions u and w belonging to $C([0, T_0]; L^\infty(D, \mathbb{R}^3))$, $t \in [0, T_0]$ to write

$$\begin{aligned}
 & | \mathcal{L}^T(u)(x, t) + \mathcal{L}^D(u)(x, t) - [\mathcal{L}^T(w)(x, t) + \mathcal{L}^D(w)(x, t)] | \\
 & \leq \frac{2}{V_\delta} \int_{D \cap B_\delta(x)} \frac{J^\delta(|y-x|)}{\delta|y-x|} | \partial_S f(\sqrt{|y-x|} S(y, x, t; u)) \\
 & \quad - \partial_S f(\sqrt{|y-x|} S(y, x, t; w)) | dy \\
 & \quad + \frac{2}{V_\delta} \int_{D \cap B_\delta(x)} \frac{J^\delta(|y-x|)}{\delta|y-x|} | H^T(u)(y, x, t) - H^T(w)(y, x, t) | dy \\
 & \quad + \frac{1}{V_\delta} \int_{D \cap B_\delta(x)} \frac{\omega^\delta(|y-x|)}{\delta^2} | \partial_\theta g(\theta(y, t; u)) - \partial_\theta g(\theta(y, t; w)) | dy \tag{29} \\
 & \quad + \frac{1}{V_\delta} \int_{D \cap B_\delta(x)} \frac{\omega^\delta(|y-x|)}{\delta^2} | H^D(u)(y, t) - H^D(w)(y, t) | dy \\
 & \quad + \frac{1}{V_\delta} \int_{D \cap B_\delta(x)} \frac{\omega^\delta(|y-x|)}{\delta^2} | \partial_\theta g(\theta(x, t; u)) - \partial_\theta g(\theta(x, t; w)) | dy \\
 & \quad + \frac{1}{V_\delta} \int_{D \cap B_\delta(x)} \frac{\omega^\delta(|y-x|)}{\delta^2} | H^D(u)(x, t) - H^D(w)(x, t) | dy.
 \end{aligned}$$

Applying (18) and (19), (20), (21), (22), (23), (24), (25), (26), (27), (28), and (29) delivers the estimate

$$\begin{aligned}
 & \| \mathcal{L}^T(u)(x, t) + \mathcal{L}^D(u)(x, t) - [\mathcal{L}^T(w)(x, t) + \mathcal{L}^D(w)(x, t)] \|_{C([0,t]; L^\infty(D, \mathbb{R}^3))} \\
 & \leq \frac{C_1 + tC_2}{\delta^2} \| u - w \|_{C([0,t]; L^\infty(D, \mathbb{R}^3))}, \tag{30}
 \end{aligned}$$

where C_1 and C_2 are constants not depending on time u or w . For $T_0 > t$, we can choose a constant $L > (C_1 + T_0 C_2) / \delta^2$ and

$$\begin{aligned}
 & \| \mathcal{L}^T(u)(x, t) + \mathcal{L}^D(u)(x, t) - [\mathcal{L}^T(w)(x, t) + \mathcal{L}^D(w)(x, t)] \|_{C([0,t]; L^\infty(D, \mathbb{R}^3))} \\
 & \leq L \| u - w \|_{C([0,t]; L^\infty(D, \mathbb{R}^3))}, \text{ for all } t \in [0, T_0]. \tag{31}
 \end{aligned}$$

This proves the Lipschitz continuity, and item (2) of the theorem is proved. Note that $u(\tau) = w(\tau)$ for all $\tau \in [0, t]$ implies $\mathcal{L}^T(u)(x, t) + \mathcal{L}^D(u)(x, t) = [\mathcal{L}^T(w)(x, t) + \mathcal{L}^D(w)(x, t)]$ and $\mathcal{L}^T(u)(x, t) + \mathcal{L}^D(u)(x, t)$ is a Volterra operator.

We write evolutions $u(x, t)$ belonging to $C([0, t]; L^\infty(D, \mathbb{R}^3))$ as $u(t)$ and $(Vu)(t)$ is the sum

$$(Vu)(t) = \mathcal{L}^T(u)(t) + \mathcal{L}^D(u)(t). \tag{32}$$

We seek the unique fixed point of $u(t)=(Iu)(t)$ where I maps $C([0, t]; L^\infty(D, \mathbb{R}^3))$ into itself and is defined by

$$(Iu)(t) = u_0 + tv_0 + \int_0^t (t - \tau)(Vu)(\tau) + b(\tau) d\tau. \tag{33}$$

This problem is equivalent to finding the unique solution of the initial value problem given by (10) and (11). We now show that I is a contraction map, and by virtue of the Banach fixed point theorem, we can assert the existence of a fixed point in $C([0, t]; L^\infty(D, \mathbb{R}^3))$. To see that I is a contraction map on $C([0, t]; L^\infty(D, \mathbb{R}^3))$, we introduce the equivalent norm

$$\| \|u\| \|_{C([0,t];L^\infty(D,\mathbb{R}^3))} = \max_{t \in [0, T_0]} \{ e^{-2LT_0 t} \|u\|_{L^\infty(D, \mathbb{R}^3)} \}, \tag{34}$$

and show I is a contraction map with respect to this norm. We apply (30) to find for $t \in [0, T_0]$ that

$$\begin{aligned} \|(Iu)(t) - (Iw)(t)\|_{L^\infty(D, \mathbb{R}^3)} &\leq \int_0^t (t - \tau) \|(Vu)(\tau) - (Vw)(\tau)\|_{L^\infty(D, \mathbb{R}^3)} d\tau \\ &\leq LT_0 \int_0^t \|u - w\|_{C([0,\tau];L^\infty(D,\mathbb{R}^3))} d\tau \\ &\leq LT_0 \int_0^t \max_{s \in [0,\tau]} \{ \|u(s) - w(s)\|_{L^\infty(D, \mathbb{R}^3)} e^{-2LT_0 s} \} e^{2LT_0 \tau} d\tau \\ &\leq \frac{e^{2LT_0 t} - 1}{2} \| \|u - w\| \|_{C([0,T_0];L^\infty(D,\mathbb{R}^3))}, \end{aligned} \tag{35}$$

and we conclude

$$\| \| (Iu)(t) - (Iw)(t) \| \|_{C([0,T_0];L^\infty(D,\mathbb{R}^3))} \leq \frac{1}{2} \| \|u - w\| \|_{C([0,T_0];L^\infty(D,\mathbb{R}^3))}, \tag{36}$$

so I is a contraction map. From the Banach fixed point theorem, there is a unique fixed point $u(t)$ belonging to $C([0, T_0]; L^\infty(D, \mathbb{R}^3))$, and it is evident from (33) that $u(t)$ also belongs to $C^2([0, T_0]; L^\infty(D, \mathbb{R}^3))$. This concludes the proof of Theorem 1.

Energy Balance

The evolution is shown to exhibit a balance of energy at all times. In this section, we describe the potential and the energy dissipation rate and show energy balance in rate form. The potential energy at time t for the evolution is denoted by $U(t)$ and is given by

$$\begin{aligned}
 U(t) = & \frac{2}{V_\delta} \int_D \int_{D \cap B_\delta(x)} \frac{J^\delta(|y-x|)}{\delta} H^T(u)(y, x, t) f(\sqrt{|y-x|} S(y, x, t; u)) dy dx \\
 & + \int_D \frac{1}{\delta^2} H^D(u)(x, t) g(\theta(x, t; u)) dx.
 \end{aligned}
 \tag{37}$$

The energy dissipation rate $\partial_t R(t)$ is

$$\begin{aligned}
 \partial_t R(t) = & -\frac{2}{V_\delta} \int_D \int_{D \cap B_\delta(x)} \frac{J^\delta(|y-x|)}{\delta} \partial_t H^T(u)(y, x, t) f(\sqrt{|y-x|} S(y, x, t; u)) dy dx \\
 & - \int_D \frac{1}{\delta^2} \partial_t H^D(u)(x, t) g(\theta(x, t; u)) dx.
 \end{aligned}
 \tag{38}$$

The derivatives $\partial_t H^T(u)(y, x, t)$ and $\partial_t H^D(u)(x, t)$ are easily seen to be nonpositive, and the dissipation rate satisfies $\partial_t R(t) \geq 0$. The kinetic energy is

$$K(t) = \rho \int_D \frac{|\partial_t u(x, t)|^2}{2} dx.
 \tag{39}$$

The energy balance in rate form is given in the following theorem.

Theorem 2. *The rate form of energy balance for the damage-fracture evolution is given by*

$$\partial_t K(t) + \partial_t U(t) + \partial_t R(t) = \int_D b(x, t) \cdot \partial_t u(x, t) dx.
 \tag{40}$$

Proof of Theorem 2. We multiply both sides of the evolution Eq. (10) by $\partial_t u(x, t)$ and integrate over D to get

$$\rho \int_D \partial_t^2 u(x, t) \cdot \partial_t u(x, t) dx = \int_D \mathcal{L}^T(u)(x, t) \cdot \partial_t u(x, t) dx
 \tag{41}$$

$$\begin{aligned}
 & + \int_D \mathcal{L}^D(u)(x, t) \cdot \partial_t u(x, t) dx \\
 & + \int_D b(x, t) \cdot \partial_t u(x, t) dx.
 \end{aligned}
 \tag{42}$$

The term on the left side of the equation is immediately recognized as $\partial_t K(t)$. The first and second terms on the right-hand side of the equation are given in the following lemma.

Lemma 3. *One has the following integration by parts formulas given by*

$$\begin{aligned} & \int_D \mathcal{L}^T(u)(x, t) \cdot \partial_t u(x, t) \, dx \\ &= -\frac{2}{V_\delta} \int_D \int_{D \cap B_\delta(x)} \frac{J^\delta(|y-x|)}{\delta} H^T(u)(y, x, t) \partial_t f(\sqrt{|y-x|} S(y, x, t; u)) \, dy \, dx. \end{aligned} \tag{43}$$

and

$$\int_D \mathcal{L}^D(u)(x, t) \cdot \partial_t u(x, t) \, dx = - \int_D \frac{1}{\delta^2} H^D(u)(x, t) \partial_t g(\theta(x, t; u)) \, dx. \tag{44}$$

Now note that

$$\begin{aligned} & \partial_t U(t) + \partial_t R(t) \\ &= \frac{2}{V_\delta} \int_D \int_{D \cap B_\delta(x)} \frac{J^\delta(|y-x|)}{\delta} H^T(u)(y, x, t) \partial_t f(\sqrt{|y-x|} S(y, x, t; u)) \, dy \, dx \\ &+ \int_D \frac{1}{\delta^2} H^D(u)(x, t) \partial_t g(\theta(x, t; u)) \, dx, \end{aligned} \tag{45}$$

and the energy balance theorem follows from (41) and (45).

We conclude by proving the integration by parts Lemma 3. We start by proving (44). We expand $\partial_t g(\theta(x, t))$

$$\begin{aligned} & \partial_t g(\theta(x, t; u)) \\ &= \partial_\theta g(\theta(x, t; u)) \frac{1}{V_\delta} \int_{D \cap B_\delta(x)} \omega^\delta(|y-x|) |y-x| \frac{\partial_t u(y) - \partial_t u(x)}{|y-x|} \cdot e_{y-x} \, dy \end{aligned} \tag{46}$$

and write

$$- \int_D \frac{1}{\delta^2} H^D(u)(x, t) \partial_t g(\theta(x, t; u)) \, dx = A(t) + B(t), \tag{47}$$

where

$$A(t) = - \int_D \frac{1}{\delta^2} H^D(u)(x, t) \partial_\theta g(\theta(x, t; u)) \frac{1}{V_\delta} \int_{D \cap B_\delta(x)} \omega^\delta(|y-x|) \partial_t u(y) \cdot e_{y-x} \, dy \, dx \tag{48}$$

and

$$B(t) = \int_D \frac{1}{\delta^2} H^D(u)(x, t) \partial_\theta g(\theta(x, t; u)) \frac{1}{V_\delta} \int_{D \cap B_\delta(x)} \omega^\delta(|y-x|) \partial_t u(x) \cdot e_{y-x} dy dx. \tag{49}$$

Next introduce the characteristic function of D denoted by $\chi_D(x)$ taking the value one inside D and zero outside, and together with the properties of $\omega^\delta(|y-x|)$, we rewrite $A(t)$ as

$$A(t) = - \int_{\mathbb{R}^3 \times \mathbb{R}^3} \chi_D(x) \chi_D(y) \omega^\delta(|y-x|) \frac{1}{\delta^2} H^D(u)(x, t) \partial_\theta g(\theta(x, t; u)) \frac{1}{V_\delta} \partial_t u(y) \cdot e_{y-x} dy dx; \tag{50}$$

we switch the order of integration and note $-e_{y-x} = e_{x-y}$ to obtain

$$A(t) = \int_D \frac{1}{V_\delta} \int_{D(x) \cap B_\delta(x)} \frac{\omega^\delta(|y-x|)}{\delta^2} H^D(u)(x, t) \partial_\theta g(\theta(x, t; u)) e_{x-y} dx \cdot \partial_t u(y) dy. \tag{51}$$

We can move $\partial_t u(x)$ outside the inner integral, regroup factors, and write $B(t)$ as

$$B(t) = \int_D \frac{1}{V_\delta} \int_{D \cap B_\delta(x)} \frac{\omega^\delta(|y-x|)}{\delta^2} H^D(u)(x, t) \partial_\theta g(\theta(x, t; u)) e_{y-x} dy \cdot \partial_t u(x) dx. \tag{52}$$

We rename the inner variable of integration y and the outer variable x in (51) and add equations (51) and (52) to get

$$A(t) + B(t) = \int_D \mathcal{L}^D(u)(x, t) \cdot \partial_t u(x, t) dx \tag{53}$$

and (44) is proved.

The steps used to prove (43) are similar to the proof of (44), so we provide only the key points of its derivation below. We expand $\partial_t f(\sqrt{|y-x|}S)$ to get

$$\begin{aligned} & \partial_t f(\sqrt{|y-x|}S(y, x, t; u)) \\ &= \partial_S f(\sqrt{|y-x|}S(y, x, t; u)) \frac{\partial_t u(y) - \partial_t u(x)}{|y-x|} \cdot e_{y-x}, \end{aligned} \tag{54}$$

and write

$$\begin{aligned} & - \frac{2}{V_\delta} \int_D \int_{D \cap B_\delta(x)} \frac{J^\delta(|y-x|)}{\delta} H^T(u)(y, x, t) \partial_t f(\sqrt{|y-x|}S(y, x, t; u)) dy dx \\ &= A(t) + B(t), \end{aligned} \tag{55}$$

where

$$\begin{aligned}
 A(t) &= \\
 &= - \int_D \frac{1}{V_\delta} \int_{D \cap B_\delta(x)} \frac{J^\delta(|y-x|)}{\delta|y-x|} H^T(u)(y, x, t) \partial_S \\
 &\quad f(\sqrt{|y-x|} S(y, x, t; u)) \partial_t u(y) \cdot e_{y-x} dy dx
 \end{aligned} \tag{56}$$

and

$$\begin{aligned}
 B(t) &= \\
 &= \int_D \frac{1}{V_\delta} \int_{D \cap B_\delta(x)} \frac{J^\delta(|y-x|)}{\delta|y-x|} H^T(u)(y, x, t) \partial_S \\
 &\quad f(\sqrt{|y-x|} S(y, x, t; u)) \partial_t u(x) \cdot e_{y-x} dy dx.
 \end{aligned} \tag{57}$$

We note that $S(y, x, t; u) = S(x, y, t; u)$ and $H^T(u)(y, x, t) = H^T(u)(x, y, t)$, and proceeding as in the proof of (44), we change the order of integration in (56) noting that $-e_{y-x} = e_{x-y}$ to get

$$\begin{aligned}
 A(t) &= \\
 &= \int_D \frac{1}{V_\delta} \int_{D \cap B_\delta(y)} \frac{J^\delta(|y-x|)}{\delta|y-x|} H^T(u)(x, y, t) \partial_S \\
 &\quad f(\sqrt{|y-x|} S(x, y, t; u)) e_{x-y} dx \cdot \partial_t u(y) dy.
 \end{aligned} \tag{58}$$

Taking $\partial_t u(x)$ outside the inner integral in (57) gives

$$\begin{aligned}
 B(t) &= \\
 &= \int_D \frac{1}{V_\delta} \int_{D \cap B_\delta(x)} \frac{J^\delta(|y-x|)}{\delta|y-x|} H^T(u)(y, x, t) \partial_S \\
 &\quad f(\sqrt{|y-x|} S(y, x, t; u)) e_{y-x} dy \cdot \partial_t u(x) dx.
 \end{aligned} \tag{59}$$

We conclude noting that now

$$A(t) + B(t) = \int_D \mathcal{L}^T(u)(x, t) \cdot \partial_t u(x, t) dx, \tag{60}$$

and (43) is proved.

Explicit Damage Models, Cyclic Loading, and Strain to Failure

In this section, we provide concrete examples of the damage functions $H^T(u)(y, x, t)$ and $H^D(u)(x, t)$. We provide an example of cyclic loading and the associated degradation in the nonlocal force-strain law as well as the strain to failure curve for monotonically increasing strains. In this work, both damage functions H^T and H^D are given in terms of the function h . Here we give an example of $h(x) : \mathbb{R} \rightarrow \mathbb{R}^+$ as follows

$$h(x) = \begin{cases} \bar{h}(x/x_c), & \text{for } x \in (0, x_c), \\ 1, & \text{for } x \leq 0, \\ 0, & \text{for } x \geq x_c. \end{cases} \tag{61}$$

with $\bar{h} : [0, 1] \rightarrow \mathbb{R}^+$ is defined as

$$\bar{h}(x) = \exp\left[1 - \frac{1}{1 - (x/x_c)^a}\right] \tag{62}$$

where $a > 1$ is fixed. Clearly, $\bar{h}(0) = 1, \bar{h}(x_c) = 0$ (see Fig. 5).

For a given critical strain $S_c > 0$, we define the threshold function for tensile strain $j_S(x)$ as follows

$$j_S(x) := \begin{cases} \bar{j}(x/S_c), & \forall x \in [S_c, \infty), \\ 0, & \text{otherwise.} \end{cases} \tag{63}$$

where $\bar{j} : [1, \infty) \rightarrow \mathbb{R}^+$ is given by

Fig. 5 Plot of $h(x)$ with $a = 2$

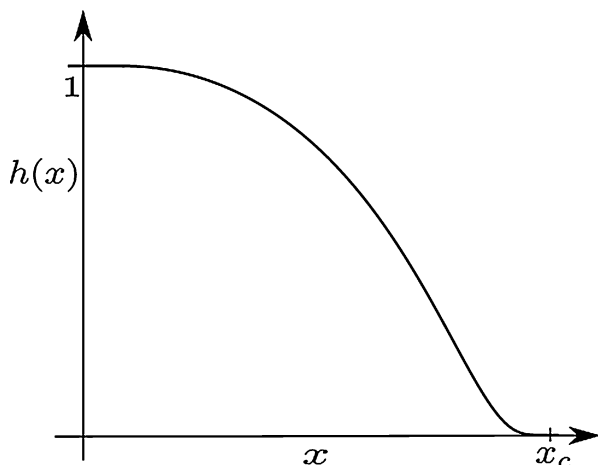
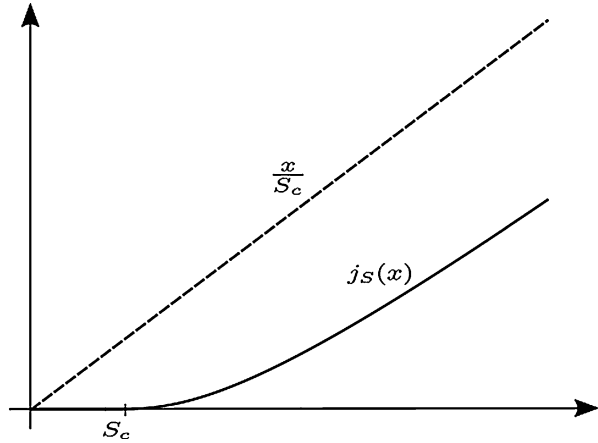


Fig. 6 Plot of $j_S(x)$ with $a = 4, b = 5$ and $S_c = 2$



$$\bar{j}(x) = \frac{(x - 1)^a}{1 + x^b} \tag{64}$$

with $a > 1$ and $b \geq a - 1$ fixed. Note that $j_S(1) = 0$. Here the condition $b \geq a - 1$ insures the existence of a constant $\gamma > 0$ for which

$$j_S(x) \leq \gamma|x|, \quad \forall x \in \mathbb{R} \tag{65}$$

(see Fig. 6).

For a given critical hydrostatic strains $\theta_c^- < 0 < \theta_c^+$, we define the threshold function $j_\theta(x)$ as

$$j_\theta(x) := \begin{cases} \bar{j}(x/\theta_c^+), & \forall x \in [\theta_c^+, \infty), \\ \bar{j}(-x/\theta_c^-), & \forall x \in (-\infty, -\theta_c^-], \\ 0, & \text{otherwise,} \end{cases} \tag{66}$$

where $\bar{j}(x)$ is defined by (64), and we plot j_θ in Fig. 7. We summarize noting that an explicit form for $H^T(u)(y, x, t)$ is obtained by using (61) and (63) in (8) and an explicit form for $H^D(u)(x, t)$ is obtained by using (61) and (66) in (9).

We first provide an example of cyclic damage incurred by a periodically varying tensile strain. Let x, y be two fixed material points with $|y - x| < \delta$, and let $S(y, x, t; u) = S(t)$ correspond to a temporally periodic strain (see Fig. 8a). Here $S(t)$ periodically takes excursions above the critical strain S_c . During the first period, we have

$$S(t) = \begin{cases} t, & \forall t \in [0, S_c + \epsilon], \\ 2(S_c + \epsilon) - t & \forall t \in (S_c + \epsilon, 2(S_c + \epsilon)] \end{cases}$$

Fig. 7 Plot of $j_\theta(x)$ with $a = 4, b = 5, \theta_c^+ = 2$, and $\theta_c^- = 3$

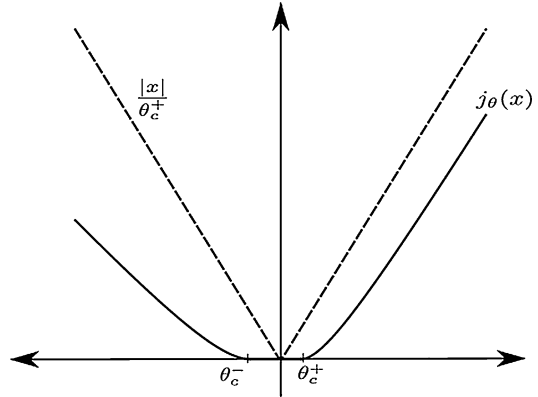
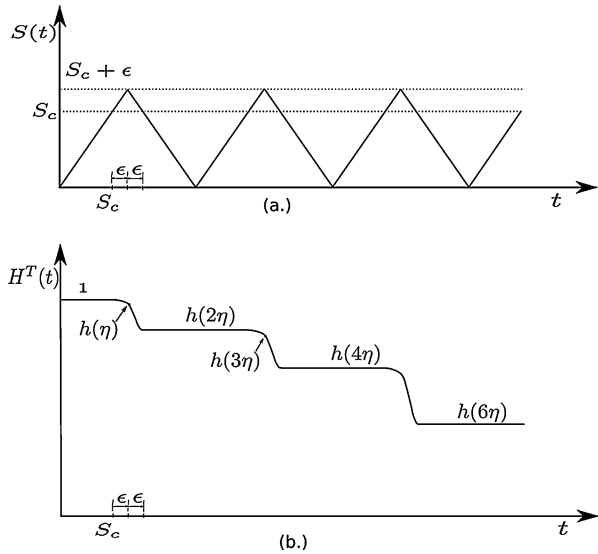


Fig. 8 (a) Strain profile. (b) Damage function plot corresponding to strain profile



and $S(t)$ is extended to \mathbb{R}^+ by periodicity (see Fig. 8a). For this damage model, we let η be the area under the curve $j_S(x)$ from $x = S_c$ to $x = S_c + \epsilon$. It is given by

$$\eta = \int_{S_c}^{S_c + \epsilon} j_S(x) dx = \int_{S_c}^{S_c + \epsilon} j_S(S(t)) dt.$$

From symmetry, the area under the curve $j_S(x)$ under unloading from $S_c + \epsilon$ to S_c is also η . The corresponding damage function $H^T(u)(y, x, t)$ is plotted in Fig. 8b.

In Fig. 9, we plot the strain-force relation where S is the abscissa and the tensile force given by $H^T(u)(y, x, t) \partial_S f(\sqrt{|y - x|} S(y, x, t; u))$ is the ordinate. Here the damage factor $H^T(u)(y, x, t)$ drops in value with each cycle of strain loading. After each cycle, the slope (elasticity) in the linear and recoverable part of the force-

Fig. 9 Cyclic strain vs Force plot. The initial stiffness is α . Hysteresis is evident in this model

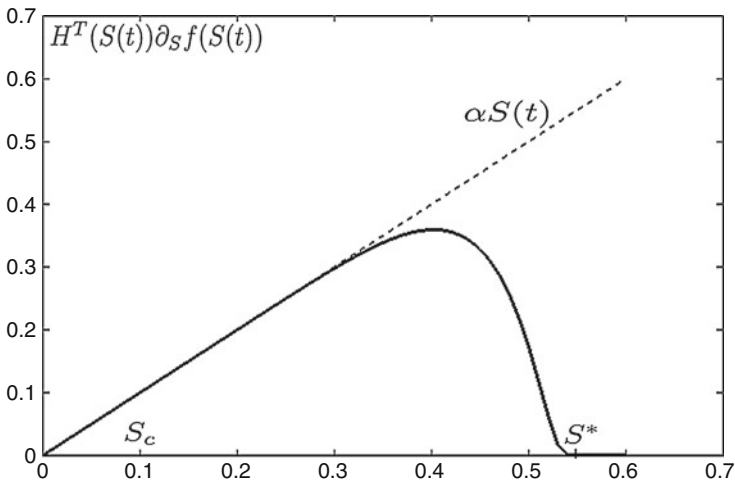
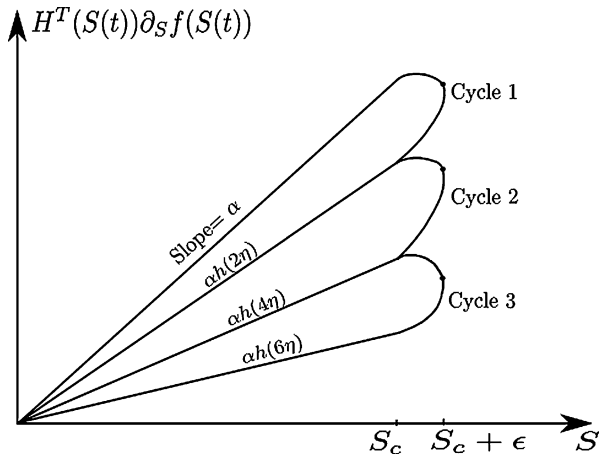


Fig. 10 Strain vs force plot where $S(t) = t$. $H^T(S(t))$ begins to drop at $S_c = 0.1$ and $S^* \approx 0.55025$

strain curve decreases due to damage. The force needed to soften the material is the strength, and it is clear from the model that the strength decreases after each cycle due to damage.

Application of this rigorously established model to fatigue is a topic of future research but beyond the scope of this article. We note that fatigue models based on peridynamic bond softening are introduced in Oterkus (2010) and with fatigue crack nucleation in the context of the Paris law in Silling and Askari (2014).

The next example is strain to failure for a monotonically increasing strain. Here we let

$S(y, x, t; u) = S(t) = t$ and plot the corresponding force-strain curve in Fig. 10.

We see that the force-strain relation is initially linear until the strain exceeds S_c ; the force then reaches its maximum and subsequently softens to failure. At $S^* \approx 0.55025$, we have $\int_0^{S^*} j_S(t)dt = x_c$ and $H^T = 0$. Here we take $\alpha = 1$.

Numerical Results

In this section, we present numerical results. Explicit expressions of the functions described in the previous section are used in simulating the problem. The damage function h is defined similar to Eq. 61 with exponents $a = 1.01$ and $x_c = 0.2$. The function j_S is given by Eq. 63 with $a = 5, b = 5, S_c = 0.01$. The function j_θ is given by Eq. 66 with $a = 4, b = 5, \theta_c^+ = 0.3, \theta_c^- = 0.4$. Nonlinear potential function f is given by $f(r) = \alpha r^2$ for $r < r_1$ and $f(r) = r$ for $r > r_2$. We let $\alpha = 10$ and let $r_1 = r_2 = 0.05$. Similarly, the nonlinear potential function g is given by $g(r) = \beta r^2$ for $r < r_1^*$ and $g(r) = r$ for $r > r_2^*$. We let $\beta = 1$ and let $r_1^* = r_2^* = 0.05$. The influence function is given by $J^\delta(|y - x|) = \omega^\delta(|y - x|) = 1 - \frac{|y-x|}{\delta}$ for $0 \leq |y - x| \leq \delta$ and $J^\delta(|y - x|) = \omega^\delta(|y - x|) = 0$ otherwise. We consider θ_c^+ and θ_c^- sufficiently high so that we only see damage due to tensile forces and not hydrostatic forces.

In both numerical problems, we consider the material domain $D = [0, 1]^2$. We also keep the initial condition fixed to $u_0 = 0$ and $v_0 = 0$. Further, we apply no body force, i.e., $b = 0$. However we will consider boundary loading that is periodic in time. Let $x = (x_1, x_2)$ where x_1 corresponds to the component along horizontal axis and x_2 corresponds to the component along vertical axis.

Periodic Loading

We apply boundary condition $u = 0$ on edge $x_1 = 0, x_1 = 1$, and $x_2 = 0$. We consider function \bar{u} of form

$$\bar{u}(t) = \begin{cases} \alpha_{bc}t, & \forall t \in [0, T_{bc}], \\ \alpha_{bc}T_{bc} - t & \forall t \in (T_{bc}, 2T_{bc}] \end{cases} \tag{67}$$

and periodically extend the function for any time t . For point x on edge $x_2 = 1$, we apply $u(t, x) = (u_1(t, x), u_2(t, x)) = (0, \bar{u}(t))$. We consider $\alpha_{bc} = 0.01$ and $T_{bc} = 0.216$.

To numerically approximate the evolution equation, we discretize the domain D uniformly with mesh size $h = \delta/5$, where $\delta = 0.15$ in this problem. For time discretization, we consider the velocity Verlet scheme for second order in time differential equation and a midpoint quadrature for the spatial discretization. Final time is $T = 1.2$ and size of time step is $\Delta t = 10^{-5}$.

To obtain the hysteresis plot, we chose bonds as shown in Fig. 11. We track the bond strain $S(y, x, t; u)$ and other relevant quantities. While we track all the bonds

Fig. 11 Discretization of material domain $D = [0, 1]^2$. During simulation bond between red and black material point is tracked to obtain the strain vs stress profile and other information

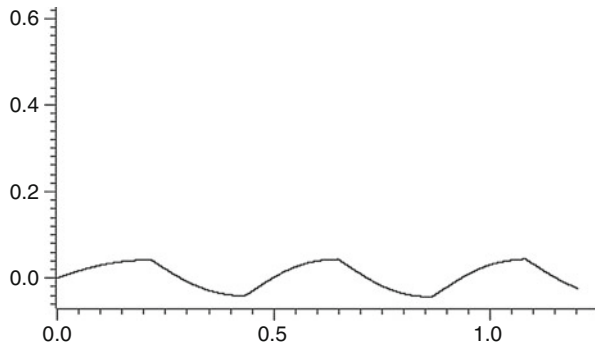
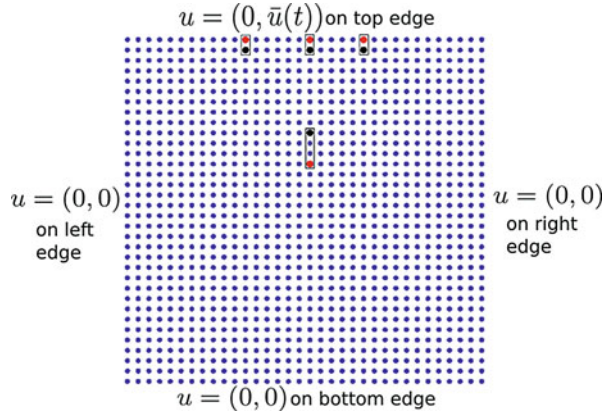


Fig. 12 Time vs Strain $S(y, x, t; u)$ plot

shown in Fig. 11, we only provide plots for the bond which is near to middle top edge. For the bonds in either left and right of the bond at middle top edge, the response is the same. For the bond inside the material, the strains are never greater than S_c , and therefore it experiences no damage.

Figures 12 and 13 show the strain of the bond and damage H^T of the bond as function of time. It is quite similar to the plots shown in Figs. 8 and 9. In Fig. 14, we show the strain vs force plot. Red line shows response of bond when damage function is taken to be unity. We further note that the damage is defined for positive strains above critical strain.

Shear Loading

We apply $u = 0$ on bottom edge and keep left and right edge free. On top, we apply $u(t, x) = (u_1(t, x), u_2(t, x)) = (\gamma t x_2, 0)$. We chose $\gamma = 0.0001$ and simulate the problem up to time $T = 750$. Time step is $\Delta t = 10^{-5}$.

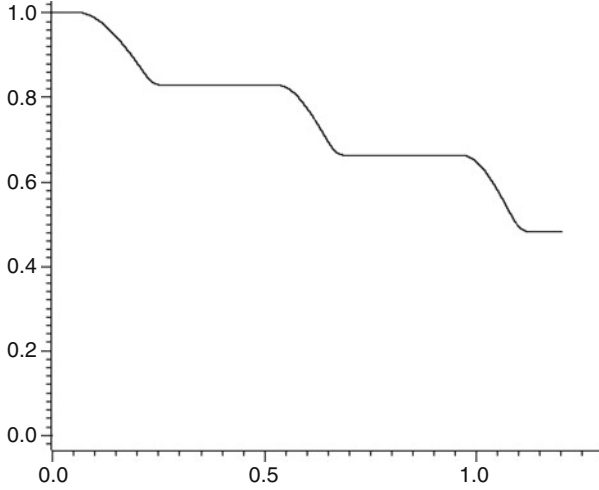


Fig. 13 Time vs Damage function $H^T((u)(y, x, t))$ plot

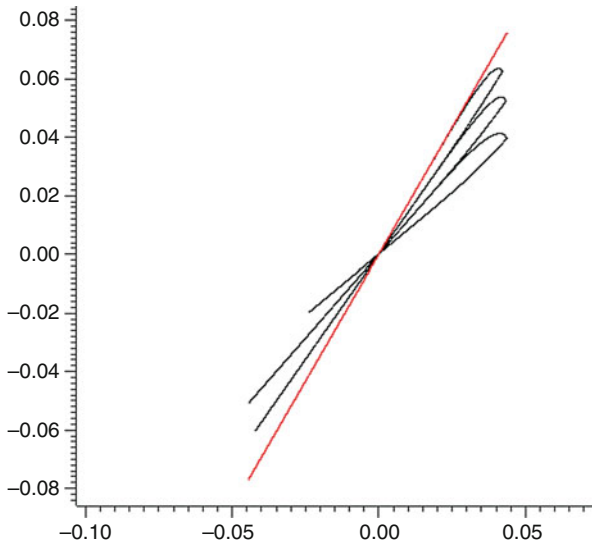
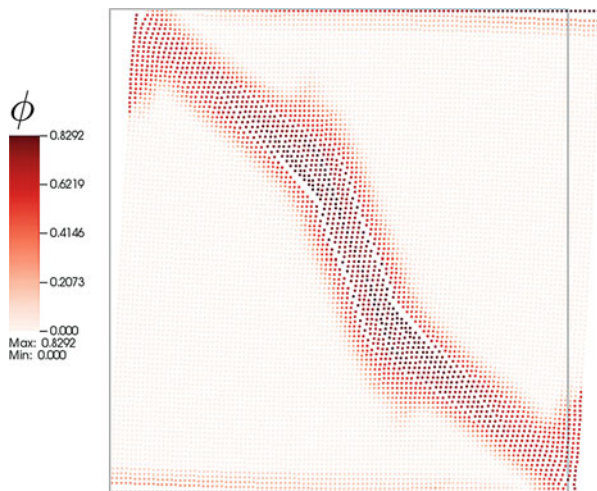


Fig. 14 Strain $S(y, x, t; u)$ vs Stress $H^T((u)(y, x, t))\partial_S f(\sqrt{|y-x|}S(y, x, t; u))$ plot for the bond near middle top edge. Red color corresponds to $\partial_S f(\sqrt{|y-x|}S(y, x, t; u))$

We choose the size of horizon to be $\delta = 0.05$ and mesh size $h = \delta/5$. As noted in the beginning of the section that we choose hydrostatic parameters large enough such that the damage is only due to the tensile interaction between material points. For tensile interaction, the extent of damage experienced by a material is defined as

Fig. 15 Each point in figure shows the discretized mesh node. Strength of color shows the damage ϕ experienced by the mesh node. Box shows reference material domain $[0, 1]^2$



$$\phi(t, x; u) = 1 - \frac{\int_{D \cap B_\delta(x)} H^T(u)(y, x, t) dy}{\int_{D \cap B_\delta(x)} dy}. \quad (68)$$

Clearly, if all bonds in a horizon of material point x suffer no damage, then ϕ will be 0. As the damage of bonds increases, ϕ also increases. In Fig. 15, we show ϕ at final time $t = 750$. As we can see, the damage is along the diagonal of square.

Linear Elastic Operators in the Small Horizon Limit

In this section, we consider smooth evolutions u in space and show that away from damage set, the operators $\mathcal{L}^T + \mathcal{L}^D$ acting on u converge to the operator of linear elasticity in the limit of vanishing nonlocality. We denote the damage set by \tilde{D} and consider any open undamaged set D' interior to D with its boundary a finite distance away from the boundary of D and the damage set \tilde{D} . In what follows, we suppose that the nonlocal horizon δ is smaller than the distance separating the boundary of D' from the boundaries of D and \tilde{D} .

Theorem 3. *Convergence to linear elastic operators. Suppose that $u(x, t) \in C^2([0, T_0], C^3(D, \mathbb{R}^3))$ and no damage, i.e., $H^T(y, x, t) = 1$ and $H^D(x, t) = 1$, for every $x \in D' \subset D \setminus \tilde{D}$, then there is a constant $C > 0$ independent of nonlocal horizon δ such that for every (x, t) in $D' \times [0, T_0]$, one has*

$$|\mathcal{L}^T(u(t)) + \mathcal{L}^D(u(t)) - \nabla \cdot \mathbb{C} \mathcal{E}(u(t))| < C\delta, \quad (69)$$

where the the elastic strain is $\mathcal{E}(u) = (\nabla u + (\nabla u)^T)/2$ and the elastic tensor is isotropic and given by

$$\mathbb{C}_{ijkl} = 2\mu \left(\frac{\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}}{2} \right) + \lambda\delta_{ij}\delta_{kl}, \tag{70}$$

with shear modulus μ and Lamé coefficient λ given by

$$\mu = \frac{f''(0)}{10} \int_0^1 r^3 J(r) dr \text{ and } \lambda = g''(0) \left(\int_0^1 r^3 J(r) dr \right)^2 + \frac{f''(0)}{10} \int_0^1 r^3 J(r) dr. \tag{71}$$

The numbers $f''(0) = \alpha$ and $g''(0) = \beta$ can be chosen independently and can be any pair of real numbers such that \mathbb{C} is positive definite.

Proof. We start by showing

$$|\mathcal{L}^T(u(t)) - \frac{f''(0)}{2\omega_3} \int_{B_1(0)} e|\xi|J(|\xi|)e_i e_j e_k d\xi \partial_{jk}^2 u_i(x)| < C\delta, \tag{72}$$

where $\omega_3 = 4\pi/3$ and $e = e_{y-x}$ are unit vectors on the sphere; here repeated indices indicate summation. To see this, recall the formula for $\mathcal{L}^T(u)$ and write $\partial_S f(\sqrt{|y-x|}S) = f'(\sqrt{|y-x|}S)\sqrt{|y-x|}$. Now Taylor expand $f'(\sqrt{|y-x|}S)$ in $\sqrt{|y-x|}S$, and Taylor expand $u(y)$ about x , denoting e_{y-x} by e to find that all odd terms in e integrate to zero and

$$\begin{aligned} |\mathcal{L}^T(u(t))_l - \frac{2}{V_\delta} \int_{B_\delta(x)} \frac{J^\delta(|y-x|)}{\delta|y-x|} \frac{f''(0)}{4} |y-x|^2 \partial_{jk}^2 u_i(x) e_i e_j e_k e_l, dy| \\ < C\delta, \quad l = 1, 2, 3. \end{aligned} \tag{73}$$

On changing variables $\xi = (y-x)/\delta$, we recover (72). Now we show

$$\begin{aligned} |\mathcal{L}^D(u(t))_k - \frac{1}{\omega_3} \int_{B_1(0)} |\xi|\omega(|\xi|)e_i e_j d\xi \frac{g''(0)}{\omega_3} \int_{B_1(0)} |\xi|\omega(|\xi|)e_k e_l d\xi \partial_{lj}^2 u_i(x)| \\ < C\delta, \quad k = 1, 2, 3. \end{aligned} \tag{74}$$

We note for $x \in D'$ that $D \cap B_\delta(x) = B_\delta(x)$ and the integrand in the second term of (6) is odd and the integral vanishes. For the first term in (6), we Taylor expand $\partial_\theta g(\theta)$ about $\theta = 0$ and Taylor expand $u(z)$ about y inside $\theta(y, t)$ noting that terms odd in $e = e_{z-y}$ integrate to zero to get

$$|\partial_\theta g(\theta(y, t)) - g''(0) \frac{1}{V_\delta} \int_{B_\delta(y)} \omega^\delta(|z-y|)|z-y|\partial_j u_i(y) e_i e_j dz| < C\delta^3. \tag{75}$$

Now substitution for the approximation to $\partial_\theta g(\theta(y, t))$ in the definition of \mathcal{L}^D gives

$$\begin{aligned}
 &|\mathcal{L}^D(u) \\
 &-\frac{1}{V_\delta} \int_{B_\delta(x)} \frac{\omega^\delta(|y-x|)}{\delta^2} e_{y-x} \frac{1}{V_\delta} \int_{B_\delta(y)} \omega^\delta(|z-y|) |z-y| g''(0) \partial_j u_i(y) e_i e_j dz dy \Big| < C \delta.
 \end{aligned} \tag{76}$$

We Taylor expand $\partial_j u_i(y)$ about x , note that odd terms involving tensor products of e_{y-x} vanish when integrated with respect to y in $B_\delta(x)$, and we obtain (74).

We now calculate as in (Lipton 2016, equation (6.64)) to find that

$$\begin{aligned}
 &\frac{f''(0)}{2\omega_3} \int_{B_1(0)} |\xi| J(|\xi|) e_i e_j e_k e_l d\xi \partial_{jk}^2 u_i(x) \\
 &= \left(2\mu_1 \left(\frac{\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}}{2} \right) + \lambda_1 \delta_{ij} \delta_{kl} \right) \partial_{jk}^2 u_i(x),
 \end{aligned} \tag{77}$$

where

$$\mu_1 = \lambda_1 = \frac{f''(0)}{10} \int_0^1 r^3 J(r) dr. \tag{78}$$

Next observe that a straight forward calculation gives

$$\frac{1}{\omega_3} \int_{B_1(0)} |\xi| \omega(|\xi|) e_i e_j d\xi = \delta_{ij} \int_0^1 r^3 \omega(r) dr, \tag{79}$$

and we deduce that

$$\begin{aligned}
 &\frac{1}{\omega_3} \int_{B_1(0)} |\xi| \omega(|\xi|) e_i e_j d\xi \frac{g''(0)}{\omega_3} \int_{B_1(0)} |\xi| \omega(|\xi|) e_k e_l d\xi \partial_{ij}^2 u_i(x) \\
 &= g''(0) \left(\int_0^1 r^3 \omega(r) dr \right)^2 \delta_{ij} \delta_{kl} \partial_{ij}^2 u_i(x).
 \end{aligned} \tag{80}$$

Theorem 3 follows on adding (77) and (80) □

Conclusions

We have introduced a simple nonlocal model for free damage propagation in solids. In this model, there is only one equation, and it describes the dynamics of the displacement using Newton’s law $F = ma$. The damage is a consequence of displacement history and diminishes the force-strain law as damage accumulates. The modeling allows for both cyclic damage or damage due to abrupt loading. The damage is irreversible, and the damage set grows with time. The dissipation

energy due to damage together with the kinetic and potential energy satisfies energy balance at every instant of the evolution. Future work will address the question of localization of damage using this model. We believe that if the loading is such that large monotonically increasing strains are generated, then damage localization based on material softening and inertia could be anticipated.

In this treatment, we have considered dynamic problems only. For this case, we have shown uniqueness for the model. The analysis of this model in the absence of inertial forces leads to the quasi-static case where the effects of inertia are absent but memory of the load history is still present. Future work aims to explore this model for this case and understand regimes of body force specimen geometry and boundary loads for which there is loss of uniqueness and associated instability. Such nonuniqueness is well known for quasi-static gradient damage models (Pham and Marigo 2013).

Acknowledgements This material is based upon the work supported by the U S Army Research Laboratory and the U S Army Research Office under contract/grant number W911NF1610456.

References

- A. Agwai, I. Guven, E. Madenci, Predicting crack propagation with peridynamics: a comparative study. *Int. J. Fract.* **171**, 65–78 (2011)
- F. Bobaru, W. Hu, The meaning, selection, and use of the peridynamic horizon and its relation to crack branching in brittle materials. *Int. J. Fract.* **176**, 215–222 (2012)
- F. Bobaru, J.T. Foster, P.H. Geubelle, S.A. Silling, *Handbook of Peridynamic Modeling* (CRC Press, BOCA Ratone, 2016)
- K. Dayal, K. Bhattacharya, Kinetics of phase transformations in the peridynamic formulation of continuum mechanics. *J. Mech. Phys. Solids* **54**, 1811–1842 (2006)
- Q. Du, Y. Tao, X. Tian, A peridynamic model of fracture mechanics with bond-breaking. *J. Elast.* (2017). <https://doi.org/10.1007/s10659-017-9661-2>
- E. Emmrich, D. Phust, A short note on modeling damage in peridynamics. *J. Elast.* **123**, 245–252 (2016)
- E. Emmrich, O. Weckner, On the well-posedness of the linear peridynamic model and its convergence towards the Navier equation of linear elasticity. *Commun. Math. Sci.* **5**, 851–864 (2007)
- J.T. Foster, S.A. Silling, W. Chen, An energy based failure criterion for use with peridynamic states. *Int. J. Multiscale Comput. Eng.* **9**, 675–688 (2011)
- W. Gerstle, N. Sau, S. Silling, Peridynamic modeling of concrete structures. *Nuclear Eng. Des.* **237**, 1250–1258 (2007)
- Y.D. Ha, F. Bobaru, Studies of dynamic crack propagation and crack branching with peridynamics. *Int. J. Fract.* **162**, 229–244 (2010)
- P. K. Jha, R. Lipton, Numerical analysis of peridynamic models in Hölder space, arXiv preprint arXiv:1701.02818 (2017)
- R. Lipton, Dynamic brittle fracture as a small horizon limit of peridynamics. *J. Elast.* **117**, 21–50 (2014)
- R. Lipton, Cohesive dynamics and brittle fracture. *J. Elast.* **124**(2), 143–191 (2016)
- R. Lipton, S. Silling, R. Lehoucq, Complex fracture nucleation and evolution with nonlocal elastodynamics. arXiv preprint arXiv:1602.00247 (2016)
- R. Lipton, E. Said, P.K. Jha, Free damage propagation with memory. *J. Elast.* To appear in *J. Elasticity* (2018)

- T. Mengesha, Q. Du, Nonlocal constrained value problems for a linear peridynamic Navier equation. *J. Elast.* **116**, 27–51 (2014)
- E. Oterkus, I. Guven, E. Madenci, Fatigue failure model with peridynamic theory, in *IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm)*, Las Vegas, June 2010, pp. 1–6
- K. Pham, J.J. Marigo, From the onset of damage to rupture: construction of responses with damage localization for a general class of gradient damage models. *Contin. Mech. Thermodyn.* **25**, 147–171 (2013)
- S.A. Silling, Reformulation of elasticity theory for discontinuities and long-range forces. *J. Mech. Phys. Solids* **48**, 175–209 (2000)
- S.A. Silling, E. Askari, A meshfree method based on the peridynamic model of solid mechanics. *Comput. Struct.* **83**, 1526–1535 (2005)
- S.A. Silling, E. Askari, Peridynamic model for fatigue cracking. Sandia Report, SAND2014-18590, 2014
- S.A. Silling, F. Bobaru, Peridynamic modeling of membranes and fibers. *Int. J. Nonlinear Mech.* **40**, 395–409 (2005)
- S.A. Silling, M. Epton, O. Weckner, J. Xu, E. Askari, Peridynamic states and constitutive modeling. *J. Elast.* **88**, 151–184 (2007)
- S.A. Silling, R.B. Lehoucq, Convergence of peridynamics to classical elasticity theory. *J. Elast.* **93**, 13–37 (2008)
- S. Silling, O. Weckner, E. Askari, F. Bobaru, Crack nucleation in a peridynamic solid. *Int. J. Fract.* **162**, 219–227 (2010)
- O. Weckner, R. Abeyaratne, The effect of long-range forces on the dynamics of a bar. *J. Mech. Phys. Solids* **53**, 705–728 (2005)
- O. Weckner, E. Emmrich, Numerical simulation of the dynamics of a nonlocal, inhomogeneous, infinite bar. *J. Comput. Appl. Mech.* **6**, 311–319 (2005)

Index

A

ABAQUS, 1111, 1112
 documentation, 1089
 simulation, 109
ABAQUS-APBS co-simulation protocol, 109
Abstract Cauchy Problem (ACP), 768, 769
Adiabatic process
 internal state variables, evolution equations
 for, 760
 microdamage nucleation and growth,
 763–764
 thermo-elastic range, 760–762
 viscoplastic range, 762
AIDD model, *see* Atomistically-informed
 interface dislocation dynamics (AIDD)
 model
Albany, 1233
Aluminium polycrystals, 679
Amorphous, 317, 329–331, 379, 380, 382, 385
Atomistic simulation, 52, 59
Atomistically-informed interface dislocation
 dynamics (AIDD) model, 163, 164
Axial vibration of micro-rods, strain gradient
 theory, *see* Micro-rods
AztecOO, 1234

B

Banach space, 1424
Bending of fractional beams, 872
Beremin model, 1102
Biaxial test, 698
Bicrystal metals, 24
Biharmonic operator, 1319
Bilinear collocation method, 1342–1346
Bivariate function, 1297, 1312
Blending, 1225
Bond breakage criterion, 1187
Bond-based linear peridynamic model,
 1335–1336

Borelian measure, 748
Boundary condition (BC)
 Cahn-Hilliard, 1319
 Dirichlet, 1295, 1299, 1302, 1306, 1308,
 1309, 1314–1317
 effects, 45
 Neumann, 1295, 1299, 1302, 1306, 1308,
 1313, 1314, 1317
 simply supported, 1319
 verification of, 1302, 1314–1316
Boundary value problem, polycrystals, 675
Brownian motion, 1332
Buckling amplitude, 261, 279–281
Buckling mode
 bending strain, 254
 circular, 254
 continuum model for helical, 245
 deformability of, 243
 helical, 243
 in-plane, 243
 out-of-plane, 249
 pre-compressed, 242
 regulation, 243
 SiNW, 268
 strain energy of, 275
Buckling wavelength, 244
 Euler beam buckling theory with, 285
 in-plane mode, 270
 of SiNW, 267, 270

C

Cahn-Hilliard (CH) equations, 1247
Cahn-Hilliard (CH) type BC, 1319
Canonical kernel function, 1296
Cantor ternary set, 911
Caputo's derivative, 859
Carpinteri column, 910
Cauchy-Green fractional deformation
 tensor, 882

- Cauchy-Green tensors, 750
- Cauchy rate dependent model, 706
- Cauchy stress, 1006, 1007, 1041
- Cauchy stress tensor, 646, 916, 919, 926
- Cauchy's inequality, 1280, 1467
- Cementitious materials, 1075–1076
 - microplane model M4, 1077–1078
 - microplane model M7, 1078–1084
- Centrosymmetry parameter (CSP), 44
- Ceramic coatings (CHCs), 155
- Chain(s), 378, 379, 391
- Chain rule, 864
- CHCs, *see* Ceramic coatings (CHCs)
- Circular crack propagation of PVB laminated glass
 - crack length and energy absorption, 490
 - crack morphology, 472
 - crack propagation, 474
 - experiment condition, 471, 472
 - impact velocity effect, 477
 - PVB thickness effect, 479
 - radial crack number and crack velocity, 485
 - statistical analysis, 482
- Circumferential cracks, 444
- Clapeyron's theorem, 930
- Clausius-Duheim inequality, 605
- Classical differentiability, 1249
- Classical Hertzian pressure calculation method, 442, 443
- Classical plasticity theory, 825
- Classical spring-slider model, 700
- Clausius-Duhem inequality, 606, 922
- CM/CS co-simulation protocol, 108
- Coherent twin boundary (CTB), 60, 63
- Cohesive strength, 1102
- Cohesive tensile force, 1268
- Cohesive-zone finite element-based model, 1365
- Cohesive zone model (CZM), 296, 305, 306, 308, 369
- Coleman-Gurtin thermodynamic procedure, 613
- Coleman-Noll procedure, 607
- Commutativity property, 1305, 1306
- Complete band-gaps, 728, 735, 737
- Composite materials, 1354
- Configuration, 748
 - deformative, 748
- Conjugate gradient (CG) method, 1384
- Constitutive equations, 647
- Constitutive laws for Cosserat continuum, 695
- Constrained crystalline strip
 - with double slips, 947
- Continuous Galerkin methods, 1210
- Continuous stiffness measurement (CSM)
 - nanindentation, strain rate, 316–318
 - experimental procedure, 321
 - indentation size effect, 326
 - indentation strain rate, 320
 - nanindentation analysis, 318
 - sample preparation, 318
 - strain rate, variation of, 322
 - variation of P/P , 322
- Continuous wave electron paramagnetic resonance (cwEPR), 80
- Continuum damage mechanics, 1106
 - modified Leckie-Hayhurst form, 1108–1109
 - plasticity driven Lemaitre form, 1109–1110
- Continuum damage model
 - constitutive equations, 520
 - microdamage, 522
- Continuum mechanics, 83–86, 99, 103, 107, 108, 124, 244, 284, 1142
- Continuum solvation approach, 99
 - Conventional continuum plasticity model, 782, 786
- Convergence, 1419, 1453–1455, 1458, 1492
 - of approximation, 1483–1488
 - rate, 1459, 1482
 - results, 1464–1465
- Correspondence model, 1189
- Cosserat continuum, 688
 - advantages, 698
 - constitutive laws for, 695
 - energy balance, 693
 - homogenization and upscaling methods, 702
 - linear and angular momentum balance, 691
 - mass balance, 692
 - phase transitions, 694
 - rock shear layer, 704
 - second law of thermodynamics, 694
- Coulomb's friction law, 450
- Coupled continuum mechanical-continuum solvation approach, 99
- Couple stress theory, 1143
- Coupling methods, LtN, *see* Local-to-nonlocal (LtN) coupling methods
- Crack-band models, 1068
- Crack path attractor, 1373
- Crack propagation
 - boundary/loading conditions, 1367
 - dynamic brittle fracture, 1369
 - path evolution, 1376
 - problem setting, 1364
- Crack-tip function, 448

- Critical strain
 dependency of, 236
 existence, 234
 variation of, 237
- CrN/Cu/Si(001) specimens, 170
- Crystal Analysis Tool, 44, 51
- Crystal plasticity
 characteristics, 941
 classical single, 597
 micropolar single, 602
- Cumulative probability, 483
- D**
- Damage model, 1497, 1512–1515, 1522
- Damage propagation
 damage evolution, existence and uniqueness of, 1501–1507
 damage models, cyclic loading and strain to failure, 1512–1515
 energy balance, 1508–1511
 formulation, 1497–1501
 periodic loading, 1516
 shear loading, 1517
 small horizon limit, linear elastic operators in, 1519–1521
- Damage stresses, 454
- Decomposition property, 1297
- Deformable grain boundary, 824
- Deformation, 378–381, 384, 385, 388, 389, 391, 397
 rates, 989, 992
 state, 1161
- Density functional theory (DFT) 158
- Detectable strain range, of indentation test, 234
- Deviatoric force state, 1182
- DFT, *see* Density functional theory (DFT)
- Dirac function, 910
- Dirichlet boundary condition, 1235, 1420
- Dirichlet data, 1232
- Dirichlet micro-boundary conditions, 612
- Discontinuous Galerkin methods, 1210
- Discrete dislocation (DD), and micropolar crystal plasticity (MP), 627, 631
- Discrete dislocation dynamics (DDD)
 simulations, 611
- Dislocation density
 distributions, 628
 tensor, 723
- Dispersion, 730
 analysis, 724–726
 curves, 727–729
 definition, 714
 in metamaterials (*see* Metamaterials)
- internal variable model, 731
 relaxed micromorphic model with curvature, 735, 736
 standard Mindlin-Eringen micromorphic model, 733
- Dissipation inequality, 1008, 1042, 1047, 1171
- Dissipative components, 555
- Dissipative constitutive equations, 613
- Dissipative gradient strengthening, 809
- Dissipative thermodynamic microforces
 grain boundary, 800
 grain interior, 794
- Double states, 1164
- Double well potentials, nonlinear nonlocal fracture models, *see* Nonlinear nonlocal fracture models
- Double well potentials, non-local, *see* Non-local double well potentials
- Drop-weight tower experiment setup, 464
- Ductile to brittle transition fracture, in ferritic steels
 continuum damage mechanics, 1106–1110
 Gurson porous plasticity, 1103–1106
 small punch testing, *see* Small punch testing (SPT)
- Ductile-to-brittle transition temperature (DBTT), 1100
- Dyadic product double state, 1164
- Dynamic brittle fracture, 1369
- Dynamic damage propagation, *see* Damage propagation
- E**
- Effective continuum medium model (ECMM), 121
- Effective stress, 1104
- Effective stress tensor, 1107
- Elastic micropolar strain, 606
- Elastic modulus, 132, 316, 319, 329, 330
- Elastic modulus ratio (EMR), 133, 135, 137, 141, 149
- Elastic property, 132
- Elastic strain, 1288
 energy, 607
 tensor, 752
- Elasto-plastic laminate microstructure, 665
- Elastoplastic properties, 217–220, 222, 225, 227
 of power-law stress-free material, 213
- Elastoviscoplasticity theory, 507
- Electro-discharge machining, 1110
- Embedded-atom method (EAM), 42, 46
- Embedded wire, 244, 275, 276
- Energetic gradient hardening, 809

- Energetic thermodynamic microforces
 grain boundary, 800
 grain interior, 794
- Energetic thermodynamic microstresses, 557
- Energy, 379–382, 384, 389, 390, 393, 394, 397
 balance, 1508–1511
 density, 1282
 dissipation rate, 1497, 1508
- Energy-blending approach, 1225
- Enriched continuum models, 716, 718–722
- Epetra*, 1234
- Equivalent microstrain model, 529
- Eringen's micromorphic framework, 645
- Eringen's nonlocal elasticity theory, 1143
- Error sensitivity, 141
- Escherichia coli*-MscL
 FEM model of, 101
 and interactions within protein, 99
- Escherichia coli*-MscS, 79
 gating and inactivation of, 83
- Euclidean point space, 749
- Euler-Almansi strain tensor, 751
- Euler-Bernoulli beam, 924, 930, 932
- Eulerian formulation, 538
- Eulerian material model, 1179
- Euler-Lagrange equation, 924, 1229
- Euler time discretization, 1458, 1463, 1465
- Euler time integration, 1052
 scheme, 1012
- Exact discretization of continuum derivatives,
 845–847
- Extended finite element method (XFEM), 448
- Extrapolated motion dynamics (EMD), 80
- F**
- Fast conjugate gradient method (FCG), 1340
- Fatigue
 definition, 402
 strength exponents, 426
 See also Indentation fatigue
- Ferritic steels, ductile to brittle transition
 fracture, *see* Ductile to brittle transition
 fracture, in ferritic steels
- Fick's approach, 1332
- Film prestress and substrate modulus,
 indentation
 fixed indenter radius, formulation for, 134
 model and computation method, 132
 variable indenter radius, general
 formulation with, 140
- Finite deformation
 equivalent microstrain model, 529
 gradient hyperelasticity, 524
 intermediate local configuration, 540
 Lagrangian strain, 531
 local objective frames, 533
 micromorphic approach, 524
 microstrain model, 526
 multiplicative decomposition, 535
 single crystal plasticity, 597
- Finite difference approximation, 1454
 consistency, 1469–1473
 convergence results, 1464–1465
 error analysis, 1465–1467
 L^2 projection, exact solution, 1467–1469
 semi-discrete approximation, 1477–1480
 stability, 1474–1477
- Finite element approximation, 1454,
 1480–1482
- Finite element implementation, 1049–1052
- Finite element method (FEM), 296, 298, 305,
 306, 308, 311, 1233
- Finite element method (FEM) simulations,
 256, 259
 ABAQUS, 279
 buckling spacing in, 260
 helical buckling configurations, 263
 membrane strain, 265
 in-plane and out-of-plane displacement
 amplitudes, 259
 in-plane displacement amplitude, 260
 in SiNW, 246
- Finite element solution procedure, 1016–1025
- Finite microstrain tensor model, 526
- First law of thermodynamics, 1166
- Fixed indenter radius, formulation
 forward analysis, 134
 reverse analysis, 135
- Flat punch indenter, 406
- Fleck-Hutchinson model, 1002
- Flow rule, 564
 grain boundary, 801
 grain interior, 795
- Fokker-Planck equation, 1332
- Force-blending, 1225
- Forward analysis, 134, 140
- Fourier heat conduction expression, 1166
- Fractal angular momentum equation, 918–920
- Fractal conservation of microinertia, 919
- Fractal continuity equation, 917
- Fractal curl operator, 913
- Fractal derivative, 912
- Fractal divergence, 913
- Fractal energy equation, 920–921
- Fractal functions, 855
- Fractal gradient, 912
- Fractal Laplacian, 913

- Fractal linear momentum equation, 918
- Fractal materials, continuum mechanics, 916–917
- Cauchy's tetrahedron, 916
 - fractal angular momentum equation, 918–920
 - fractal continuity equation, 917
 - fractal energy equation, 920–921
 - fractal linear momentum equation, 917
 - fractal second law of thermodynamics, 921–923
 - fractal wave equations, 923–926
- Fractal second law of thermodynamics, 921–923
- Fractal wave equations
- 3D elastodynamics, 925, 926
 - fractal Timoshenko beam, elastodynamics of, 923–925
- Fractional calculus
- applications, 870
 - arc length, 865
 - balance principles, 889
 - chain rule, 864
 - deformation, 884
 - deformation geometry, 881
 - fractional curvature vector, 867
 - fractional radius of curvature, 867
 - fundamental forms, 875
 - in material deformations, 857
 - in mechanics, 855
 - normal fractional curvature, 877
 - origin, 854
 - polar decomposition, 883
 - properties, 858
 - Serret-Frenet equations, 868
 - stresses, 888
 - tangent space, 866
 - vector field theorems, 879
 - vector operators, 878
 - Zener viscoelastic model, 893
- Fractional nonlocal continuum mechanics and microstructural models
- exact discretization of continuum derivatives, 845–847
 - lattice derivatives of integer orders, 843
 - lattice fractional integro-differentiation, 842–843
 - lattice models to continuum models, 844–845
 - long-range interactions of lattice particles, 841–842
- Fracture strength, 301, 312
- Fracture toughness, 374
- Free energy, 1008, 1010
- Free energy density function, 1181
- Free fracture model, non-local double well potentials, *see* Non-local double well potentials
- Free micro-inertia, 720, 721, 728, 729, 731, 733, 735, 736
- Free volume, 380, 381, 392
- Freund's theory, 462
- Fréchet derivatives, 1164
- Fubini's theorem, 931, 1312
- Fully-homogenized peridynamic (FH-PD) model, 1360
- Functional calculus (FC), 1318–1328
- Functionally graded materials (FGMs), 1354
- convergence studies, 1369
 - fully-homogenized peridynamic model, 1360
 - intermediate-homogenization peridynamic model, 1361
- G**
- Gain-of-function (GOF) mutant, 117
- Galerkin finite element method, 1334, 1338–1340
- Galerkin methods, 1210
- Gating mechanism, 82, 83, 110
- Gauss divergence theorem, 881
- Gauss theorem, 911, 1011
- Gaussian elimination (Gauss), 1340
- Generalized convolution operators, 1325
- Geomaterials, 1069
- Geometrically necessary dislocation (GND), 40, 50, 54, 56, 644, 783, 942, 1003, 1005
- density tensor, 600
 - gradient, 808
 - hardening, 812
- Glasses, 378, 380, 383, 387, 397
- Glassy polymers, *see* Continuous stiffness measurement (CSM) nanoindentation experiments
- Glassy polymers, *see* Plastic deformation
- Gradient-enhanced nonlocal plasticity theory, 783
- Gradient free energy, 1014
- Gradient micro-inertia, 720, 721, 724, 727–733, 735, 736
- Grain boundary (GB), 5, 12, 14, 16, 20, 24
- deformable boundary condition, 824
 - dissipative thermodynamic microforces, 799
 - effects, 40, 59
 - energetic thermodynamic microforces, 800

- Grain boundary (GB) (*cont.*)
 flow rule, 801
 microfree boundary condition, 814
 microhard boundary condition, 821
 model and SGCP (*see* Strain gradient
 crystal plasticity (SGCP))
 simulation model, 371
 virtual power, 798
- Green-Gauss theorem, 926
- Green-Lagrange strain tensor, 750
- Green's theorem, 879
- Green strain tensors, 438
- Griffith fracture energy, 1458
- Griffith free energy, 1454
- Griffith's theory, 928
- Gronwall's inequality, 1278
- Gurson porous plasticity, 1103–1106
- Gurson-Tvergaard-Needleman (GTN) model,
 1101, 1102, 1113
- H**
- Hölder inequality, 1478
- Hölder space
 cohesive dynamics, finite time
 intervals, 1425
 Lipschitz continuity and bound, 1424–1425
 local existence and uniqueness, 1433–1436
- Hölder's inequality, 1259
- Hall-Petch effect, 1036
- Hardness, 316–320, 322, 326, 330, 331
- Heat flow, 796
- HEC, *see* Hydrogen embrittlement cracking
 (HEC)
- Helical mode, 243, 245, 247, 254, 267, 271,
 272, 276, 281–284
- Helmholtz decomposition, 914
- Helmholtz free energy, 557
- Helmholtz free energy density, 648
- Hertzian cone crack, 444
- High-strength steels, 290, 291
- Higher differentiability, 1256–1261
- Higher integrability, 1255–1256
- Higher-order beam theory, 1125
- Higher order gradient plasticity
 dissipative thermodynamic
 microstresses, 561
 energetic and dissipative components, 554
 flow rule, 564
 Helmholtz free energy and energetic
 thermodynamic microstresses, 557
 thermodynamic formulation, 553
 thermo-mechanical coupled heat
 equation, 565
- Higher order nonlocal problems, 1261–1262
- Higher-order rod model, micro-rods, *see*
 Micro-rods
- Homogeneous deformations, 885
- Homogenization of fractal media, 914–915
 anisotropic fractals, vector calculus on,
 913–914
 fractional integral theorems and fractal
 derivatives, 911–913
 mass power law, 907–908
 product measure, 907–911
- Hooke law, 926
- Hoop stress, 472
- Hourglass control, 701
- Hydrogen embrittlement cracking (HEC),
 290, 291
 crack length and stress intensity
 factor, 311
 crack propagation, mechanism of, 305
 experimental results, 303
 materials and experimental
 methods, 302
 multiple crack formation, mechanism
 of, 306
See also Threshold stress intensity factor,
 HE cracking
- Hydrostatic energy release rate, 1287
- Hydrostatic force, 1268
- Hydrostatic softening, 1277
- Hydrostatic strain, 1267, 1268, 1286,
 1498, 1513
- I**
- ICME, *see* Integrated computational materials
 engineering (ICME)
- Iffpack*, 1234
- Influence function, 1188
- Impact, 1357, 1364, 1368, 1373, 1376
- Incremental theory, 783
- Indentation, 169, 403
 depth, 418, 425
 image of polycrystalline copper, 419
 instrumented, 212
 load-depth curve, 404
 load-displacement curve, 213
 load-displacement curve vs. material
 property, 214
 response, 213
 size effect, 5, 23, 31, 35
 shape factors, 225
 spherical analysis, 213, 423
 strain rate, 317, 318, 320–322, 329
 stress intensity factor, 406

- Indentation fatigue, 402
 damage, 422
 deformation, 409
 depth, 424
 depth propagation law, 409
 mechanics theory of, 404
 numerical simulation of, 411
 on polycrystalline copper, 415
 on semi-infinite solid, 405
 strength law, 422
 steady-state rate of, 413, 416
 testing system, 414
- Indentation, prestressed elastic
 coating/substrate system, 130
 fixed indenter radius, formulation for, 134
 model and computation method, 132
 variable indenter radius, general
 formulation with, 140
- Indentation test, HEC, *see* Hydrogen
 embrittlement cracking (HEC)
- Indenter angle, 214, 215, 218, 219, 223, 225,
 227–229
- Indistinguishable load-displacement curve, 239
- Infinitesimal deformations, 888
- Initial Boundary Value Problem (IBVP),
 765–769
- In-plane mode, 254, 267, 268, 270, 271,
 283, 284
- Integral operator, 1307
- Integrated computational materials engineering
 (ICME), 161
- Interaction domain, 1202
- Interfaces
 dynamic properties, 367
 elastic constants, 366
 fracture, 369
 mechanical properties, 363
- Intermediate-homogenization peridynamic
 model, 1361
- Internal state variable, 609
- Internal variable model, 729–731
- Isotropy, 1168
- J**
- Jensen's inequality, 931, 1467, 1469
- Jumarie's derivatives, 860
- K**
- Kernel(s), 1248–1249
- Kernel function, 1296, 1298, 1300, 1301, 1303,
 1305, 1307, 1312, 1316, 1326
- Khun-Tucker conditions, 985
- Kinematic(s), 786
- Kinematically admissible displacement
 field, 927
- Kirchhoff stress tensor, 438, 755
- L**
- Lévy process, 1332
- Lagrangian formulation, 536
- Lagrangian virtual work method, 438
- Lamé coefficient, 1520
- Laplace equation, 1254
- Laplace operator, 1319, 1326
- Laser spallation test, 157
- Lattice antiderivative, 844
- Lattices fractional integro-differentiation,
 842–843
- Lebesgue dominated convergence theorem,
 1296, 1306, 1309, 1314
- Leckie-Hayhurst form, 1108–1109
- Leibniz L-fractional derivative, 861
- Leibniz rule, 1296, 1306, 1314
- Length scale, nanoindentation,
see Nanoindentation, size effects
 and material length scales 4
- Length scale parameter, 1124, 1133, 1137,
 1142–1145, 1152, 1153
- Lennard–Jones (LJ) potential, 42, 46
- Lennard–Jones (LJ) type, 1267
- Linear peridynamic solid (LPS) model,
 1230, 1391
- Linear regression analysis, 484
- Linearized kinematics, 601
- Lipid bilayers, continuum modeling of, 104
- Lipid membrane model, equi-biaxial tension
 of, 105
- Lipschitz continuity, 1475, 1503, 1506
- Lipschitz property, 1486, 1487
- Loading curvature, 215–219, 222, 231, 232,
 234, 235
- Loading rate, 317, 320, 321, 331
- Local Laplacian, 1249–1254
- Local-to-nonlocal (LtN) coupling methods
 for continuum mechanics, 1225–1226
 discretization of, 1233–1234
 linearized linear peridynamic solid and
 classical elasticity, 1231–1232
- M**
- Macroscopic boundary conditions, 823
- Macroscopic stress, 976
- Major failure on-set (MFO) stress and
 strain, 459

- Mandel stress tensor, 648
 Manifold, 748
 Material body, 748
 Material continuum, 748
 Material point, 748
 Matrix-vector multiplication, 1333
 MDN, *see* Misfit dislocation network (MDN)
 MEAM, *see* Modified embedded atom method (MEAM)
 Mechanosensitive (MS) channels, 79, 119
 Mechanosensitive channel of small/large conductance (MscS/MscL), 79
 Mechanotransduction, 82
 Median vent crack, 444
 Meshfree method, 1233
 Metal/ceramic interfacial regions (MCIRs)
 ab-initio density functional theory, 158
 AIDD simulator and VPSC-CIDD model, 163
 CrN/Cu/Si(001) specimens, 170
 elastic stiffness constants, 177
 glancing incidence XRD pattern, 171, 172
 grain structure by TEM, 175
 hardness, 178
 ICME, 161
 indentation, 177, 178
 interfacial impurities/solute-atoms, 165
 interfacial spacing, 160
 laser spallation test, 157
 MD simulations, 159
 micro-and nano-scale roughness, 166
 micro-pillar fabrication and axial compression measurements, 180
 microscale mechanical testing, 159
 misfit dislocation network, 159
 multiscale simulation, 166
 nano adhesion interlayers, 155
 nanolaminate adhesion interlayer, 165
 nanolaminate composites, 155
 selected area diffraction patterns, 176
 shear failure, 167
 TEM, 176
 tensile stress, 159, 160
 vapor phase deposition, 167
 Young's modulus, 180
See also Ti/TiN interfaces/Ti/TiN interfaces
 Metal matrix composite, 633
 Metamaterials, 715, 736
 band-gap metamaterials, 716–719
 dynamics of, 721
 mechanical behavior of, 720
 periodic microstructures, 716
 specific inertial characteristics of, 720
 Micro-clamped boundary conditions, 948
 Micro-free boundary conditions, 947
 Micromodulus double state, 1171
 Micromorphic theory
 constitutive equations, 510
 elastic–plastic decomposition, 517
 elastoviscoplasticity theory, 507
 non-dissipative mechanisms, 513
 Micro-Electro Mechanical Systems (MEMS), 805
 Micromorphic crystal plasticity free energy potentials for, 662
 Micro-pillar fabrication
 axial compression testing, 180
 CrN/Cu/Si vs. CrN/Ti/Si pillars, 185
 cylindrical micro-pillars, 182
 ex-situ and in-situ compression testing, 186
 in-situ axial compression, 183
 in-situ compression testing, 185
 interfacial locking and unlocking, 189
 NanoFlip device, 180
 stage of deformation, 189
 using scripted FIB milling, 180
 Micro-rods, 1142, 1153
 axial vibration problem, solution of, 1148
 clamped-free and clamped-clamped, 1152
 equation of motion, 1147
 geometry and coordinate system, 1145
 Hamilton's principle, 1147
 initial conditions and boundary conditions, 1148
 kinetic energy, variation of, 1147
 modified strain gradient theory, formulation for, 1144
 non-zero strain component, 1145
 strain energy, variation of, 1146
 symmetric rotation gradient tensor, 1146
 Microbeam, 1124
 free vibration, 1132
 modified strain gradient elasticity theory, 1125
 trigonometric shear deformation beam theory, 1127
 Microcurl model, 658, 661
 Microdamage model, 502
 Microdeformation, 648
 Microfree boundary condition, 815
 Microhard boundary condition, 822
 Micromorphic crystal plasticity, micropolar and, 651
 Micromorphic model with curvature, 734–736
 Micromorphic theory
 elastic–plastic decomposition, 508
 kinematics of, 504

- Microplane models, 1067, 1069–1070
 cementitious materials, 1075–1084
 kinematic no-split formulation, 1070–1072
 kinematic split formulation, 1072–1074
 nonlocal microplane models, 1085–1095
 static split formulation, 1074–1075
- Micropolar crystal plasticity (MP), discrete dislocation (DD) and, 627
- Micropolar single crystal simulations, 621
- Micropolar theory, 617
- Microscopic boundary condition, 823
- Microscopic stress, 976
- Microstress vectors, 1006, 1008
- Microstructure, 378, 382, 393, 715, 736
 characteristic sizes of, 715
 inertia of, 721
 periodic microstructures, 716
 vibrations of, 716, 721
- Mid-point integration scheme, 1384
- Misfit dislocation network (MDN), 159, 193
- Model M7Auto, 1086–1095
- Modified couple stress theory, 1143
- Modified embedded atom method (MEAM), 42, 166
- Modified Leckie-Hayhurst form, 1108–1109
- Modified shear correction factor, 1133
- Modified strain gradient elasticity theory, 1125
- Modified strain gradient theory (MSGT)
 formulation for, 1144
 natural longitudinal frequencies, 1152
 results of, 1153
- Molecular dynamics (MD) simulation, 40, 370
 boundary conditions effects, 45
 grain boundary effects, 59
 simulation methodology, 41
 small length scales, size effects in, 52
 theoretical models, 50
- Molecular dynamics-decorated finite element method (MDeFEM), 82
- Momentum theorem, 479
- Morse potential, 42
- MTS 810 Hydraulic Materials Testing System, 954
- Multi-scale simulation, 83, 85
- Mühlhaus-Vardoulakis Cosserat plasticity model, 706
- N**
- Nano adhesion interlayers, 155
- Nanocrystalline structure, 834
- NanoFlip device, 180
- Nanoindentation, *see* Continuous stiffness measurement (CSM)
- Nanoindentation, molecular dynamics,
see Molecular dynamics (MD)
 simulation
- Nanoindentation, size effects and material length scales
 bicrystal metals, grain boundaries on, 24
 length scale, determination of, 11
 nonlocal theory, 6
 physically based material length scale, 8
 sample preparations, 16
 single crystal and polycrystalline metals,
 temperature and strain rate dependency
 on, 18
- Nano/micro-electromechanical systems (N/MEMS), 402
- Nanowire, 243, 244
- Natural frequency, 1152, 1153
- Navier-Cauchy equation (NCE), 1231
- Navier solution technique, 1137
- N -dimensional vectors, 1343
- Neumann-type conditions, 1203
- Newton's second law, 1420
- Newton-Raphson iterations, 1017
 scheme, 988
- Newton-Raphson solution procedure, 1052
 scheme, 1012
- Non-convex bond based material, 1177
- Non-homogeneous deformation, 887
- Non-incremental theory, 783
- Non-linear Poisson-Boltzmann (NLPB) model, 107
- Non-local double well potentials
 energy release rate, 1286–1288
 existence of solutions, 1270–1273
 linear elastic operators, 1288–1290
 nonlocal dynamics, 1267–1270
 peridynamic energy to elastic properties,
 1281–1286
 softening zone, control of, 1276–1281
 stability analysis, 1273–1276
- Non-local effects, 724, 731, 736
- Non-proportional loading, 550
- Nonlinear nonlocal fracture models
 H^2 norm, Lipschitz continuity in,
 1437–1453
 H^2 solution, existence of, 1437–1453
 Hölder continuous solution, existence of,
 1425–1436
 Hölder norm, Lipschitz continuity in,
 1425–1436
 Hölder space, solutions in, 1423–1425
 nonlocal potential, 1420–1422
 Sobolev space H^2 , solutions in, 1436–1437
 weak formulation, 1422

- Nonlocal Bernoulli-Euler beam model, 1143
- Nonlocal crystal plasticity (NCP)
 model, 941
 strain gradient formulation, 943
- Nonlocal diffusion model, 1334–1337,
 1342–1346
- Nonlocal force, 1428
- Nonlocal Green's identities, 1201
- Nonlocal interactions, 1278, 1290, 1496
- Nonlocal Laplacian
 definition, 1244
 elliptic properties for, 1249–1250
 operator, 1202
 scaling, 1250–1254
- Nonlocal microplane models, 1085–1095
- Nonlocal model(s), 1224, 1460–1461
 accuracy of, 1224
 finite difference approximation, *see* Finite
 difference approximation
 finite element approximation, *see* Finite
 element approximation
 and local, *see* Local-to-nonlocal (LtN)
 coupling methods
 higher differentiability, 1256–1261
 higher integrability, 1255–1256
 higher-order, 1246–1247
 peridynamics equation, 1461–1463
 regularity for, 1261–1262
- Nonlocal operators, 1245, 1249
- Nonlocal operators, local boundary conditions
 in 1D, 1302–1311
 in 2D, 1296–1302, 1311–1314
 higher dimensions, operators in,
 1316–1317
 numerical experiments, 1317–1324
 unctional calculus, 1318–1328
 verification of BC, 1314–1316
- Nonlocal theory, 6
- Nonlocal wave equation, 1317–1321,
 1324, 1328
- Nooru-Mohamed mixed mode fracture
 test, 1089
- Numerical discretization, 1384
- Numerical study, 213, 239
- O**
- Optimization-based coupling (OBC) methods
 domain configuration, 1227
 goal of, 1227
 LtN formulation, 1231–1234
 non-intrusive, 1226
 patch tests, 1234–1235
 rectangular bar, crack, 1235–1238
- state models and properties, 1229–1231
- tensile test specimen, crack, 1238–1240
- well-posedness, 1227–1229
- P**
- Pairwise bond force density field, 1165
- Palmitoyloleoylphosphatidylethanolamine
 (POPE) lipid bilayer, 104
- Partial differential equation (PDE), 1200, 1226,
 1234, 1244, 1245, 1332
- Partial integro-differential equations (PIDEs),
 1244, 1246
- Partial-volume algorithms, 1396
- Patch tests, 1234–1235
- Peridigm*, 1234
- Peridynamics (PD), 1224, 1227, 1229,
 1231–1232, 1234, 1235, 1244, 1294,
 1295, 1326, 1328, 1356, 1359, 1378
 discretization schemes, 1205–1208
 equation of motion, 1199
 local limits, 1200
 multiscale finite element implementation,
 1210–1212
 nonlocal obstacle problems, 1217
 nonlocal operators, 1201
 nonlocal optimal control problems,
 1212–1214
 nonlocal parameter identification problems,
 1214, 1217
 nonlocal vector calculus, 1200
 steady-state problem, 1202, 1203, 1205
 variational formulation, 1204
- Peridynamic energy, 1477
- Peridynamic equilibrium equation, 1230
- Peridynamic model, 1334–1337, 1418,
 1459, 1488
 fast collocation method, 1346–1349
 and meshfree discretization, 1391
 one-dimensional, 1337–1342
- Peridynamic problem
 dynamic, convergence studies of, 1405
 static, convergence studies
 of, 1398
- Peridynamic theory, 1332
 damage state, 1173
 balance of momentum, 1164
 bond based viscoelastic material, 1176
 bond-based linear material with
 damage, 1174
 concepts in, 1161
 convergence to local theory, 1190
 damage evolution, 1183
 discrete systems, 1179

- energy balance and thermodynamics, 1166
 - energy density, 1278
 - force, 1420, 1421, 1437, 1439, 1486, 1487
 - heat transport model, restriction on, 1171
 - isotropic bond based material, 1177
 - local continuum damage mechanics, 1193
 - local kinematics and kinetics, 1188
 - material model, 1167
 - non-convex bond based material, 1177
 - non-ordinary state based material, 1176
 - ordinary state based material with
 - damage, 1175
 - plasticity, 1180
 - properties of states, 1163
 - purpose of, 1160
 - thermodynamic formulation of material models, 1169
 - Permanent deformation state, 1181
 - Perturbation analysis, 1273
 - Phase field model, 516
 - Phase velocity, 715
 - Phonon damping mechanism, 747
 - Piola–Kirchhoff stress tensor, 648, 926
 - Plastic cracks, 444
 - Plastic deformation, glassy polymers
 - free volume evolution, 381
 - rate, effect of, 389
 - shear banding and indentation size effect, 395
 - STZ nucleation energy evolution, 384
 - temperature, effect of, 391
 - thermal history, 393
 - Plastic energy density, 989, 994
 - Plastic flow, 1181
 - Plasticity, 380–382
 - Plastic slip, 1004–1013
 - Plastic strain rate, 1005, 1039
 - Plausible damage growth law, 1183
 - Poincaré inequality, 1247, 1248, 1250, 1254, 1259
 - Point mass, 910
 - Poisson's ratio, 132, 988, 1285
 - Polar decomposition, 750
 - Polycarbonate (PC), 317, 318, 321–324, 326, 330, 331
 - Polycrystals, 834
 - grain size effects in, 674
 - Polymeric glasses, *see* Continuous stiffness measurement (CSM) nanoindentation experiments
 - Polymers, 378–380, 382, 383, 385, 388, 389, 391, 396
 - Polymethyl methacrylate (PMMA), 317, 318, 321–323, 326, 328, 331, 468
 - Porous elastic materials, 1380
 - Post-buckling behaviors, 281
 - of embedded wires, 244, 275
 - theoretical and FEM evolutions of, 283
 - Potential energy, 1269, 1420, 1462
 - Power-law type of interactions, 841, 842
 - Pre-fractals, 906
 - Principle of virtual power, 550
 - Pulsed electron-electron double resonance (PELDOR) approach, 80
 - PVB laminated glass
 - circumferential crack propagation characteristics, 457
 - computational method, 450
 - constitutive relation, 435
 - contact model, 442
 - dynamic out-of-plane loading, 462
 - impact velocity, 446
 - internal stress analysis, 444
 - material model, 442
 - model setup, XFEM, 449
 - Poisson's ratio, 445
 - quasi-static loading, 458
 - radial crack propagation, 453
 - XFEM, 448
 - See also* Radial crack propagation of PVB laminated glass
 - radial crack propagation of PVB laminated glass
- Q**
- Quasi Newton-Raphson method, 803
 - Quasi-static fracture, 1380
- R**
- Radial crack propagation of PVB laminated glass, 453
 - crack morphology, 465
 - driving mechanisms, 469
 - drop weight and height, 470
 - experimental condition, 463
 - specimen preparation, 465
 - See also* Circular crack propagation of PVB laminated glass
 - Raman spectroscopy, 364
 - Rank-one defect energy, 665
 - Rate, 378–381, 385, 387–389, 393
 - Rate variational formulation, 1013–1016
 - Rate-dependent (RD) model, 707, 988–991, 994, 995, 997
 - analytical solutions, 983–984
 - constitutive assumptions, 979
 - equilibrium, 979–980
 - evolution, 980–981

- Rate-independent (RI) model, 988, 990, 995, 997
 analytical solutions, 984–987
 constitutive assumptions, 981
 equilibrium, 981–982
 evolution, 982
 Rayleigh wave, 1373
 Rayleigh wave speed of glass sheet, 468
 Relative deformation tensor, 603
 Relaxation time, 437
 Relaxed micromorphic model, 723–724, 731, 732, 734, 737
 and band-gap mechanical metamaterials, 718–719
 dispersion analysis, 724–726
 dispersion curves, 727–729
 with curvature, 734, 736
 Representative volume element (RVE), 757, 830, 909, 918, 1106
 Residual stress, 290, 291, 306, 311, 312
 Response curves, 989, 990, 992, 995
 Reverse analysis, 135, 141, 214, 232–234, 236, 239
 Reynolds transport theorem, 911, 912, 917, 918
 Reynold’s transfer theorem, 891
 Riemann-Liouville (R-L) derivatives, 859
 Riemannian space, 749
 Rotation axial vector, 651
 Rousselier model, 1102
- S**
- Scanning imaging with electron-or ion-induced secondary electrons (SE/ISE), 168
 Schmid law, 649
 Schmid stress, 1016
 Second law of thermodynamics, 694
 Serret-Frenet equations, 868
 Shear deformation, 1125, 1127
 Shear locking, 701
 Shear modulus, 1285, 1520
 Shear stress-strain response, 623
 Shear transformation, 380, 381, 384, 385, 387, 397
 Shear transformation zones (STZs), 380, 381, 384
 Simply supported (SS) type BC, 1319
 Size dependency, 1124, 1142, 1145, 1153
 Size effects, 782, 785, 823, 824, 832
 Small punch testing (SPT), 1118
 crack depth, effects of, 1116–1118
 experimental setup, 1110–1112
 geometric parameters of, 1115
 numerical simulations of, 1111–1113
 puncher radius, effects of, 1114–1117
 Small-scale effect, 1124, 1143
 Sobolev embedding property, 1436, 1444, 1446
 Sobolev space H^2 , 1436–1437
 Soft substrate, 275, 276
 Space, Euclidean, 748
 Spatial discretization, 1470–1473
 Split Hopkinson pressure bar (SHPB)
 compression, 492
 Stability, 1474–1477
 analysis, 1273–1276
 linearized peridynamics, 1488–1492
 semi-discrete approximation, 1477–1482
 Standard Galerkin approach, 1049
 Standard Mindlin-Eringen model, 731–734
 State based peridynamics, 1266, 1270, 1496, 1501
 Static classical elasticity, 1231
 Statically admissible field, 927
 Statistically stored dislocations (SSDs), 783, 830, 942
 Stokes’ theorem, 880
 Strain, 378, 379, 381–391, 393, 394, 396
 energy density function, 1162
 vs. force plot, 1515
 gages, 957
 profile, 1514
 vs. stress, 1518
 vs. time, 1517
 Strain gradient crystal plasticity (SGCP), 1038
 bi-crystal specimen, single slip system, 1053–1057
 bulk material, free energy imbalance, 1042–1044
 cylindrical specimen, three slip system, 1058–1059
 finite element implementation, 1049–1052
 finite element solution procedure, 1016–1025
 interface, free energy imbalance, 1045–1049
 macroscopic and microscopic energy balances, 1039–1042
 plastic slip, 1004–1013
 polycrystalline behavior, 1025–1029
 rate variational formulation of, 1013–1016
 Strain gradient plasticity (SGP), 4, 7, 35, 37, 659, 974–975
 analytical solutions, 983–987
 deformable grain boundary, 824
 dissipation, 977
 equilibrium, 977
 evolution, 978

- microfree boundary condition, 815
 - microhard boundary condition, 822
 - plastic slip patterning, 974, 988–993
 - polycrystals, 832
 - problem statement, 975
 - rate-dependent model (*see* Rate-dependent (RD) model)
 - rate-independent model (*see* Rate-independent (RI) model)
 - tensile steel bars, necking in, 992–997
 - theories, 942
 - two-dimensional finite element analysis, 804
 - validation of, 807
 - Strain gradient plasticity model
 - experimental validation, 569
 - finite element model, 566
 - Strain gradient theory, micro-rods, strain gradient theory, *see* Micro-rods
 - Strain rate, *see* Continuous stiffness measurement (CSM) nanoindentation experiments
 - Strain tensor, 1005
 - Stress, 379, 380, 382, 388–394
 - equilibrium, 1067
 - jump, 575
 - strain response for epoxy interface, 368
 - Stress intensity factor analysis (SIF), 446
 - Stress–strain curves, 459, 962
 - Stretch-passivation problem, 574
 - Substrate system, indentation, *see* Indentation, prestressed elastic coating/substrate system
 - Surface free energy, 367
 - Symmetric tensor, 1283
- T**
- Taylor-Quinney coefficient, 1106
 - Tchebychev's inequality, 1278
 - Temperature distribution, 814
 - Tensile softening, 1277
 - Tensile strain, 1267, 1269, 1277, 1512
 - Tensile toughness, 1278
 - Theory of Thermo-Viscoplasticity (TTV), 744
 - constitutive postulates, 754
 - constitutive relations, 754–756
 - crucial aspects, 744
 - definition, 747–748
 - disadvantage, 744
 - IBVP, 765–769
 - kinematics, 749–753
 - microdamage tensor, 756–759
 - Thermal softening, 704
 - Thermo-mechanical coupled heat equation, 565
 - Thermo-mechanical coupling, 784, 795, 834
 - Thermodynamic microstress
 - dissipative, 561
 - energetic, 557
 - Thermodynamics heat, flow, 796
 - Thermomechanics, 512
 - Thin film, indentation, *see* Indentation, prestressed elastic coating/substrate system
 - Threshold stress intensity factor, HE cracking, 291
 - cohesive zone model, 296
 - crack growth resistance, 300
 - experimental results, 292
 - finite element method, 298
 - materials and experimental methods, 291
 - Ti/TiN interfaces, 190
 - coherent regions and misfit dislocation network, 191
 - intrinsic energetic characteristics, 197
 - MD/DFT simulations, 190
 - misfit dislocation network location, 193
 - preferred orientation relation, 190
 - shear response, 199
 - Time discretization, 1463, 1469
 - convergence of approximation, 1483–1488
 - convergence results, 1464–1465
 - error analysis, 1465–1469
 - linearized peridynamics, stability condition for, 1488–1492
 - Timoshenko beam theory (TBT), 1125
 - Timoshenko theory of beams, 691
 - Toeplitz matrix, 1339
 - Traction-separation law (TSL), 370
 - Transition, 244, 266, 268, 273, 274, 282–284
 - Transmission electron microscopy (TEM), 1045
 - Transverse strain, 955, 957
 - Trigonometric beam model, 1125, 1127
 - Trilinos*, 1234
 - Tsai-Sun model, 369
 - Two-dimensional finite element analysis, 804
 - Two-parameter Weibull model, 483
 - Two-phase single crystal laminate, 653
- U**
- Uniform mass, 910
 - Unique solution, 213, 223, 239
 - Univariate function, 1297, 1304

V

Validation, strain gradient plasticity theory, 807
Vapor phase deposition, 167
Variable indenter radius, formulation
 error sensitivity, 141
 forward analysis, 140
 reverse analysis, 141
Vector-valued states, 1163
Vehicle speed, 446
Velocity-Verlet method, 1384
Vibration, 1132
Virtual power
 grain boundary, 798
 grain interior, 788
 method, 506
Viscoelastic stress tensors, 439
Viscoplasticity
 adiabatic process, material functions for,
 760–764
 HSLA-65 steel, material parameters
 identification for, 770, 770
 IBVP, 765–769
 relaxation time, 746–747
 spall fracture phenomenon modelling,
 771–775
 thermo-viscoplasticity model, 747–759
Visco-plastic self-consistent continuum
 interface dislocation dynamics
 (VPSC-CIDD) model, 163, 164

Viscous generalized stress, 515
Volume fraction porosity, 759
Voronoi tessellation, 1025
VPSC-CIDD model, *see* Visco-plastic
 self-consistent continuum interface
 dislocation dynamics (VPSC-CIDD)
 model

W

Weakly integrable kernels, 1245
Web cracks, 444
Weibull parameters, 483
Weierstrass function, 872
Weibull distribution, 1101
Wild type (WT) MscL, 117

X

X-ray diffraction (XRD), 168
X-ray photoelectron spectroscopy
 (XPS), 169

Y

Young's inequality, 1260
Young's modulus, 180