Alain Combescure
René de Borst
Ted Belytschko
Editors

IUTAM Bookseries

IUTAM Symposium on Discretization Methods for Evolving Discontinuities

Proceedings of the IUTAM Symposium held in Lyon, France, 4–7 September, 2006

Springer

IUTAM SYMPOSIUM ON DISCRETIZATION METHODS FOR EVOLVING DISCONTINUITIES

IUTAM BOOKSERIES

Volume 5

Aims and Scope of the Series

The IUTAM Bookseries publishes the proceedings of IUTAM symposia under the auspices of the IUTAM Board.

For a list of books published in this series, see final pages.

IUTAM Symposium on Discretization Methods for Evolving Discontinuities

Proceedings of the IUTAM Symposium held in Lyon, France, September 4–7, 2006

Edited by

ALAIN COMBESCURE

LaMCos, INSA de Lyon, France

RENÉ DE BORST

Eindhoven University of Technology, Eindhoven, The Netherlands

and

TED BELYTSCHKO

Northwestern University, Evanston, IL, USA

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN 978-1-4020-6529-3 (HB)
ISBN 978-1-4020-6530-9 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

www.springer.com

Printed on acid-free paper

Table of Contents

Partition-of-Unity Based Finite Element Methods

Other Discretization Methods

Preface

Discontinuities are an important domain in the mechanics of solids and fluids. With mechanics focusing on smaller and smaller length scales in order to understand the physics that underly many phenomena that hitherto were modelled in a phenomenological manner, the need to properly model discontinuities increases rapidly. Classical examples are cracks, shear bands and rock faults at a macroscopic level. However, the increase in computational power has made it possible to also analyse phenomena like delamination and debonding in composites (mesoscopic level) and phase boundaries and dislocation movements at the microscopic and nanoscopic level. While the above examples all pertain to solid mechanics, albeit at a wide range of scales, technically important (moving) fluid-solid interfaces appear in welding and casting processes and in aeroelasticity.

Standard discretization methods such as finite element, finite difference or boundary element methods have been developed for continuous media and are less well suited for treating evolving discontinuities. Indeed, they are approximation methods for the solution of the partial differential equations, which are valid on a domain. Discontinuities divide this domain into two or more parts and at the interface special solution methods must be employed. This holds a fortiori for moving discontinuities such as Lüders–Piobert bands, Portevin-le-Chatelier bands, solid-state phase boundaries, fluid-solid interfaces and dislocations.

In recent years, discretization methods have been proposed, which are more flexible than standard finite element methods and which have the potential of capturing (moving) discontinuities in a robust and efficient manner. For this reason it was timely to organize a IUTAM symposium that addresses these developments. This book collects 24 contributions at the symposium written by renowned scientists. Together they constitute the state-of-the-art in the field.

Alain Combescure René de Borst Ted Belytschko

MESHLESS FINITE ELEMENT METHODS

Meshless discretisation of nonlocal damage theories

Harm Askes[1], Terry Bennett[1] and Sivakumar Kulasegaram[2]

[1] University of Sheffield, Department of Civil and Structural Engineering, Sir Frederick Mappin Building, Mappin Street, Sheffield S1 3JD, United Kingdom, h.askes@sheffield.ac.uk, t.bennett@sheffield.ac.uk

[2] Cardiff University, School of Engineering, Queen's Buildings, The Parade, Cardiff CF24 0YF, United Kingdom, KulasegaramS@cf.ac.uk

Summary. In this chapter, meshless discretisation methods are explored in the implementation of nonlocal continuum damage theories. Integral-type and gradient-type nonlocality are both considered. The main advantage of using a meshless implementation (compared to more established discretisation methods such as the finite element method) is that the higher-order continuity requirements imposed by gradient-type nonlocality can be accomodated straightforwardly. Thus, meshless methods are particularly suited as an implementational framework to test and compare various nonlocal theories. Here, the element-free Galerkin (EFG) method is used. In particular, second-order and fourth order gradient damage models are compared to integral-type damage models whereby the integral nonlocal operator acts on the equivalent strain or on the displacements. No signficant differences in response are found, which implies that the inclusion of a fourth-order term in the gradient-type nonlocality is of lesser importance. Finally, the mathematical nonlocality of EFG interpolation functions is tested to ascertain whether it provides a mechanical nonlocality to the description. It is shown that this is not the case. However, despite this lack of intrinsic mechanical nonlocality, the EFG method is an excellent tool for the numerical implementation of a nonlocal continuum theory.

Key words: nonlocal damage, gradient-enhanced damage, higher-order continuum, meshless methods, continuity of interpolation

1 Introduction

Fracture of heterogeneous materials is a complicated process that often involves a number of stages. Initially, several micro-cracks are formed, distribution of which can be dispersed over a zone with significant dimensions. After the peak load is reached and beyond, the micro-cracks tend to coalesce into a macroscopically observable crack. At this stage of the loading process, the

Alain Combescure et al. (eds.), IUTAM Symposium on Discretization Methods for Evolving Discontinuities, 3–20.

deformations concentrate within this macro-crack whereas the rest of the specimen unloads. Whereas a discontinuous (either in terms of displacements or strains) representation is appropriate in the later loading stages, the inelastic processes can be modelled adequately within a continuum framework in the earlier loading stages. A realistic modelling of the entire fracture process could well involve the transition from a continuum approach to a discontinuous approach. It is therefore of interest to explore the use of advanced discretisation methods in the use of continuum material models as well as discontinuous material models. The focus of the present chapter is the use of *meshless methods* in the simulation of fracture by means of *continuum theories*.

It has been well documented that the straightforward use of standard continuum models such as plasticity and damage in the post-peak loading regime leads to the loss of mathematical well-posedness and a severe dependence on the applied discretisation in numerical simulations — see for instance the overview presented by Sluys [38] and references cited therein. Typical manifestations of these problems include a zero width of the zone in which failure takes place and a zero energy consumption, neither of which is realistic. As an alternative to standard continuum models, *nonlocality* can be included in the material model in order to avoid these deficiencies. Nonlocality assumes that an interaction exists between material particles at finite distance. By means of these interactions it is guaranteed that the failure zone has a finite width and that failure can only take place at finite energy dissipation.

Two main formats of nonlocality exist: gradient-type and integral-type nonlocality. In gradient nonlocal continuum models the field equations are equipped with additional spatial derivatives of one or more state variables. Integral nonlocality is invoked by performing spatial averaging on certain state variables. Robust gradient-enriched plasticity theories have been proposed by Aifantis [1, 2] and later been extended towards numerical implementations [14, 27, 28]. Gradient-enriched damage theories are more recent [12, 18, 30]. As regards integral-type nonlocality, nonlocal plasticity theories were explored by Eringen [16, 17] and a recent overview is provided by Jirásek and Rolshoven [23]. Nonlocal damage was suggested by Pijaudier-Cabot and Bažant [7, 34] with many follow-up works in more recent years.

The two types of nonlocality are strongly related — gradient theories can be derived from integral theories by applying Taylor series expansions as has been demonstrated on various occasions [19, 27, 30]. An appealing feature of most nonlocal continuum theories is that the width of the zone in which further failure takes place normally *decreases* whilst failure progresses, and in many cases this width tends to zero. As a result, a gradual transition from failure within a continuum framework towards discontinuous failure is obtained. This is a suitable simulation of experimental observations, in which diffuse microcracking is followed by the formation of a macroscopic crack. However, the numerical implementation of nonlocal theories is not always straightforward:

- In gradient-type nonlocal theories the maximum order of derivation in the field equations sets the continuity requirements on the interpolation functions. Whereas the standard continuum theories of mechanics are normally of second order and hence require $\mathcal{C}^0$-continuous shape functions, gradient theories are often of the order 4 or even higher, by which $\mathcal{C}^1$-continuity or higher is required.
- In integral-type nonlocal theories the spatial averaging is done by means of computing integrals over the nonlocal interaction domains. For a certain integration point these domains are not necessarily bounded by the adjacent nodes, but instead often extend over multiple layers of surrounding integration points. For numerical implementations this implies that a nonlocal connectivity array must be set up that is different from the usual element topology (which contains the connectivity between nodes and elements). To compute the tangent stiffness matrix the common element-by-element assembly of finite element packages cannot be used anymore but must be extended to account for the nonlocal interaction [15, 22, 35, 36].

Meshless methods are a class of discretisation methods that address higher-order continuity as well as nonlocal connectivity in a natural manner. Meshless methods are straightforwardly formulated with an arbitrary order of continuity, which has been exploited in various applications of linear and nonlinear gradient theories requiring $\mathcal{C}^1$-continuity [3, 4, 29, 39] or even $\mathcal{C}^2$-continuity [5,6]. Furthermore, the nonlocal connectivity array of integral nonlocal has the same structure as the node-to-integration point assembly connectivity of meshless methods, a fact which can also be exploited in numerical implementations. However, it is emphasized here that the two types of connectivity are of a different origin: the nonlocal connectivity is a *mechanical property* that appears in the continuum equations, whereas the meshless assembly connectivity is a *mathematical property* of the applied discretisation and interpolation functions. They should not be confused, and terms such as "the intrinsic nonlocality of meshless methods" are to be understood as mathematical nonlocality, not as mechanical nonlocality.

In the remainder of this chapter, meshless methods are applied to nonlocal damage theories. In section 2 the formulation of a particular type of meshless shape functions is treated, namely those of the element-free Galerkin (EFG) method. Section 3 deals with nonlocal damage theories of the integral type and the gradient type. The relation between the various gradient theories is explained, and the higher-order boundary conditions are discussed as well. Sections 4 and 5 present examples. In section 4 the various nonlocal formulations are compared, including a recently suggested integral formulation based on nonlocal displacements [21, 36, 37]. The intrinsic nonlocality of meshless interpolations is assessed in section 5 in terms of their capacity to objectively describe failure.

2 Element-free Galerkin shape functions

As common in meshless methods, element-free Galerkin (EFG) shape functions are formulated in terms of a set of nodes without elements. Instead of employing elements to set a connectivity between nodes, each node is assigned a so-called domain of influence. Inside this domain of influence, a weight function corresponding to the associated node is nonzero, while it is zero outside the domain of influence [8, 9, 25].

The approximant function u^h is assumed to be the inner product of a monomial base vector $\mathbf{p}$ and a vector with coefficients $\mathbf{a}$:

$$u^h = \mathbf{p}^T(\mathbf{x})\mathbf{a}(\mathbf{x}) \tag{1}$$

where for instance a complete quadratic monomial base vector in two dimensions is given by

$$\mathbf{p}^T(\mathbf{x}) = [1, x_1, x_2, x_1^2, x_1x_2, x_2^2] \tag{2}$$

The (yet unknown) coefficients $\mathbf{a}$ are resolved by minimising the moving least squares sum J given by

$$J = \sum_{i=1}^{n} w_i(\mathbf{x}) \left(\mathbf{p}^T(\mathbf{x}_i)\mathbf{a}(\mathbf{x}) - u_i\right)^2 \tag{3}$$

where n is the number of nodes, w_i is the weight function of node i, and u_i is the discrete parameter of node i. Minimisation of Equation (3) with respect to $\mathbf{a}$ yields

$$\left(\mathbf{P}^T\mathbf{W}(\mathbf{x})\mathbf{P}\right)\mathbf{a}(\mathbf{x}) = \left(\mathbf{P}^T\mathbf{W}(\mathbf{x})\right)\mathbf{u} \tag{4}$$

where $\mathbf{u}$ contains u_i and

$$\mathbf{P}^T = [\mathbf{p}(\mathbf{x}_1), \mathbf{p}(\mathbf{x}_2), \ldots, \mathbf{p}(\mathbf{x}_n)] \tag{5}$$

$$\mathbf{W}(\mathbf{x}) = \mathrm{diag}\,[w_1(\mathbf{x}), w_2(\mathbf{x}), \ldots, w_n(\mathbf{x})] \tag{6}$$

From Eq. (4) the coefficients $\mathbf{a}(\mathbf{x})$ can be obtained, which are then substituted into Eq. (1). Then, EFG shape functions $\mathbf{H}$ can be derived as

$$u^h = \mathbf{p}^T(\mathbf{x})\mathbf{a}(\mathbf{x}) = \mathbf{p}^T(\mathbf{x})\left(\mathbf{P}^T\mathbf{W}(\mathbf{x})\mathbf{P}\right)^{-1}\left(\mathbf{P}^T\mathbf{W}(\mathbf{x})\right)\mathbf{u} = \mathbf{H}^T(x)\mathbf{u} \tag{7}$$

The two main ingredients of the EFG shape functions are the monomial base vector $\mathbf{p}$ (or, more precisely, the terms that are included in $\mathbf{p}$) and the weight functions w. The governing parameter in the weight function is the size of the domain of influence d as compared to the nodal spacing h. In this study a circular domain of influence with radius d is taken as [9]

$$w(s) = \begin{cases} \dfrac{\exp\left(-\dfrac{s^2}{(\alpha d)^2}\right) - \exp\left(-\dfrac{d^2}{(\alpha d)^2}\right)}{1 - \exp\left(-\dfrac{d^2}{(\alpha d)^2}\right)} & \text{if } s \le d \\ 0 & \text{if } s > d \end{cases} \tag{8}$$

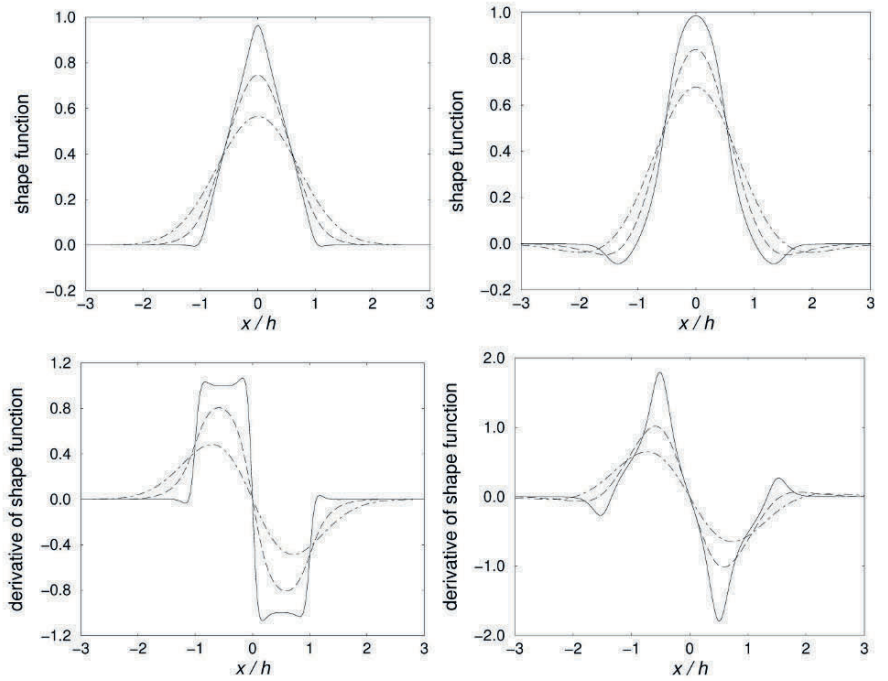


Fig. 1. EFG shape functions (top) and their spatial derivatives (bottom). Left column: $\mathbf{p}^T(\mathbf{x}) = [1, x]$ with $d/h = 2$ (solid), $d/h = 3$ (dashed) and $d/h = 4$ (dash-dotted). Right column: $\mathbf{p}^T(\mathbf{x}) = [1, x, x^2]$ with $d/h = 3$ (solid), $d/h = 4$ (dashed) and $d/h = 5$ (dash-dotted).

with $s = |\mathbf{x} - \mathbf{x}_i|$ and $\alpha = 1/4$ a numerical parameter to set the relative weights inside the domain of influence. The weight function $w(s)$ decays with increasing distance to the associated node, and it is non-negative in the entire domain of influence.

In order to carry out the matrix inversion included in Eq. (7) the domain of influence must contain a sufficient number of nodes — this is usually quantified by the parameter d/h where h is the nodal spacing. Figure 1 shows one-dimensional EFG shape functions and its spatial derivatives for linear and quadratic base vectors $\mathbf{p}$ and for a range of d/h. For relatively low values of d/h the EFG shape functions behave as finite element shape functions, whereas smoother EFG shape functions are obtained for larger values of d/h. Comprehensive works on the EFG method and its implementation include [8,9,25]. In [3,4,29] the discretisation of gradient-enriched continuum theories is treated in detail.

3 Nonlocal continuum damage theory

Isotropic continuum damage theories usually relate the stresses $\boldsymbol{\sigma}$ to the strains $\boldsymbol{\varepsilon}$ via

$$\boldsymbol{\sigma} = (1-\omega)\,\mathbf{C}\boldsymbol{\varepsilon} \tag{9}$$

where $\mathbf{C}$ contains the elastic constants and ω is the damage parameter that ranges from 0 initially to 1 upon complete failure of the material. The evolution of this damage parameter is driven by the following properties:

Damage driving state variable: A state variable must be selected that should relevant for the studied the failure process. Often, this is a strain invariant. It could be defined with different sensitivities to tensile and compressive loading, e.g. In this study, the definition according to Mazars and Pijaudier-Cabot [26] is employed:

$$\varepsilon_{\text{eq}} = \sqrt{\sum_i <\varepsilon_i>^2} \tag{10}$$

where $<\varepsilon_i> = \frac{1}{2}\varepsilon_i + \frac{1}{2}|\varepsilon_i|$ and ε_i is the i^{th} principal strain.

Loading-unloading conditions: The damage parameter ω is required to be non-decreasing. A history parameter κ is used to keep track of previous values of ε_{eq} by defining $\kappa = \max\{\varepsilon_{\text{eq}}, \kappa, \kappa_0\}$ where κ_0 is a user-defined initial value that sets the elastic limit. The history parameter κ indicates the highest level of strain (measured through ε_{eq}) that has been experienced by the material point under consideration. Loading and unloading are formally controlled by requiring that

$$\varepsilon_{\text{eq}} - \kappa \leq 0, \qquad \dot{\kappa} \geq 0, \qquad (\varepsilon_{\text{eq}} - \kappa)\dot{\kappa} = 0 \tag{11}$$

Softening law: Upon fulfilment of the conditions in Eq. (11) an evolution function for the damage variable as a function of the history variable is needed. A bi-linear uni-axial stress-strain relation, useful for theoretical observations, is obtained by taking

$$\omega = \frac{\kappa_u(\kappa - \kappa_0)}{\kappa(\kappa_u - \kappa_0)} \tag{12}$$

with κ_u another user-defined parameter that sets the uni-axial strain level at which all load-carrying capacity is exhausted.

The above continuum damage framework is still local. An elegant way to include nonlocality is through the equivalent strain definition [30,34]. Within an integral-type nonlocal framework, a nonlocal equivalent strain $\varepsilon_{\text{eq}}^{nl}$ can be cast as

$$\varepsilon_{\text{eq}}^{nl}(\mathbf{x}) = \frac{\int\limits_V \exp\left(-\frac{\boldsymbol{\xi}^T\boldsymbol{\xi}}{2\ell^2}\right)\varepsilon_{\text{eq}}(\mathbf{x}+\boldsymbol{\xi})\,\mathrm{d}V}{\int\limits_V \exp\left(-\frac{\boldsymbol{\xi}^T\boldsymbol{\xi}}{2\ell^2}\right)\,\mathrm{d}V} \tag{13}$$

which carries out an averaging of the (local) equivalent strain ε_{eq}. The averaging is weighted by the (decaying) exponential function in which the parameter ℓ sets the size of the zone over which the averaging takes place. This parameter is linked to the micro-structural dimensions and is often called the *internal length* or *intrinsic length*.

Note the similarities between the role of the nonlocal internal length ℓ in Eq. (13) and the role of the domain of influence radius d in Eq. (8): both set an interaction over finite distance (mathematical interaction in case of d, mechanical interaction in case of ℓ) and both are based on exponential functions that decay radially. The neighbour graph that links an integration point with its neighbouring integration points in Eq. (13) can be determined with the same algorithm that computes the neighbour graph linking nodes and integration points for Eq. (8).

Gradient-type nonlocality can be derived independently but it is instructive to emphasize the relation to the integral-type nonlocality. To this end, the factor $\varepsilon_{\text{eq}}(\mathbf{x}+\boldsymbol{\xi})$ in Eq. (13) can be expanded in Taylor series around $\mathbf{x}$. After some straightforward algebra [19, 27, 30], the gradient-type counterpart of Eq. (13) is found as

$$\varepsilon_{\text{eq}}^{nl} = \varepsilon_{\text{eq}} + \frac{1}{2}\ell^2\nabla^2\varepsilon_{\text{eq}} + \frac{1}{8}\ell^4\nabla^4\varepsilon_{\text{eq}} + \dots \tag{14}$$

which sets an *explicit* dependence of the nonlocal equivalent strain on the local equivalent strain. Eq. (14) can also be written as [4, 30]

$$\varepsilon_{\text{eq}}^{nl} - \frac{1}{2}\ell^2\nabla^2\varepsilon_{\text{eq}}^{nl} + \frac{1}{8}\ell^4\nabla^4\varepsilon_{\text{eq}}^{nl} + \dots = \varepsilon_{\text{eq}} \tag{15}$$

by which the dependence of $\varepsilon_{\text{eq}}^{nl}$ on ε_{eq} has become *implicit*. Eq. (15) can be derived from Eq. (14) as follows:

- Take the Laplacian ∇^2 of Eq. (14) and multiply with $\frac{1}{2}\ell^2$. Substract the result from Eq. (14), by which

$$\varepsilon_{\text{eq}}^{nl} - \frac{1}{2}\ell^2\nabla^2\varepsilon_{\text{eq}}^{nl} = \varepsilon_{\text{eq}} - \frac{1}{8}\ell^4\nabla^4\varepsilon_{\text{eq}} + \dots \tag{16}$$

- Take the squared Laplacian ∇^4 of Eq. (14) and multiply with $\frac{1}{8}\ell^4$. Add the result to Eq. (16), by which

$$\varepsilon_{\text{eq}}^{nl} - \frac{1}{2}\ell^2\nabla^2\varepsilon_{\text{eq}}^{nl} + \frac{1}{8}\ell^4\nabla^4\varepsilon_{\text{eq}}^{nl} = \varepsilon_{\text{eq}} + \dots \tag{17}$$

- This process could be repeated for every derivative of ε_{eq} that is to be replaced by a derivative of $\varepsilon_{\text{eq}}^{nl}$.

Since Eq. (15) is a differential equation in terms of the nonlocal equivalent strain, in a boundary value problem it must be accompanied by appropriate

boundary conditions. To illustrate these, the weak form of Eq. (15) is written as

$$\int_{\Omega} \delta e \left(\varepsilon_{\rm eq}^{nl} - \frac{1}{2}\nabla^2 \ell^2 \varepsilon_{\rm eq}^{nl} + \frac{1}{8}\ell^4 \nabla^4 \varepsilon_{\rm eq}^{nl} - \varepsilon_{\rm eq} \right) {\rm d}\Omega = 0 \tag{18}$$

where a truncation after the ℓ^4 term is made and δe is a test function. Integration by parts of the higher-order terms (once for the ℓ^2 term, twice for the ℓ^4 term) yields

$$\int_{\Omega} \left(\delta e\, \varepsilon_{\rm eq}^{nl} + \frac{1}{2}\ell^2 (\nabla \delta e)^T \cdot \nabla \varepsilon_{\rm eq}^{nl} + \frac{1}{8}\ell^4 \nabla^2 \delta e\, \nabla^2 \varepsilon_{\rm eq}^{nl} - \delta e\, \varepsilon_{\rm eq} \right) {\rm d}\Omega =$$
$$\oint_{\Gamma} \delta e\, \mathbf{n}^T \cdot \nabla \left(\frac{1}{2}\ell^2 \varepsilon_{\rm eq}^{nl} - \frac{1}{8}\ell^4 \nabla^2 \varepsilon_{\rm eq}^{nl} \right) {\rm d}\Gamma + \oint_{\Gamma} \mathbf{n}^T \cdot \nabla \delta e\, \frac{1}{8}\ell^4 \nabla^2 \varepsilon_{\rm eq}^{nl}\, {\rm d}\Gamma \tag{19}$$

where $\mathbf{n}$ is the normal to the boundary Γ of the domain Ω. Formally, the test function gradient $\nabla \delta e$ must be decomposed into a normal component and a tangential component, since it is impossible to independently prescribe $\varepsilon_{\rm eq}^{nl}$ together with its tangental gradient. However, in the examples below all boundaries are aligned with the Cartesian axes, by which this decomposition becomes irrelevant. In the numerical implementation, the boundary integrals have been ignored, which amounts to imposing the following natural boundary conditions:

$$\mathbf{n}^T \cdot \nabla \left(\frac{1}{2}\ell^2 \varepsilon_{\rm eq}^{nl} - \frac{1}{8}\ell^4 \nabla^2 \varepsilon_{\rm eq}^{nl} \right) = 0 \qquad \text{on } \Gamma \tag{20}$$

$$\mathbf{n}^T \frac{1}{8}\ell^4 \nabla^2 \varepsilon_{\rm eq}^{nl} = 0 \qquad \text{on } \Gamma \tag{21}$$

The boundary conditions of the second-order implicit gradient model (truncated after the ℓ^2 term) are fairly well accepted and render good results. The boundary conditions of the fourth-order implicit gradient model are more difficult to assess, and a more in-depth study would be needed should this model become more widely used.

An interesting line of recent research concerns the thermodynamic foundations of nonlocal damage theories. For instance, a thermodynamically consistent variant of the earlier framework of Pijaudier-Cabot and Bažant [34] is presented by Borino [10] for integral-type nonlocal damage; similarly, Peerlings et al. [33] give a thermodynamically consistent re-formulation of the gradient-type nonlocal damage model presented in [30]. Finally, it is noted that the nonlocality can also be formulated to operate on another variable, such as for instance the history parameter κ [13] or the damage ω itself [12, 24]. An excellent in-depth comparison between several approaches and their capacity to model failure objectively has been presented by Jirásek [20].

4 Comparison between different nonlocal formulations

The correspondence or otherwise between the various nonlocal damage formulations has been studied on various occasions [4,5,31,32]. A general conclusion is that the implicit series of Eq. (15) are to be preferred over the explicit series of Eq. (14). Furthermore, good agreement between the integral model and the second-order (up to order ℓ^2) and the fourth-order (up to order ℓ^4) truncations of Eq. (15) was found.

4.1 Wave propagation in a one-dimensional bar

The correspondence between the second-order and fourth-order implicit models according to Eq. (15) is verified numerically by means of a one-dimensional wave propagation problem. Whereas the second-order model only requires $\mathcal{C}^0$-continuity of the discretisation, $\mathcal{C}^1$-continuity is needed for the fourth-order model. However, this poses no difficulties within an EFG framework.

Figure 2 shows the boundary conditions and the dimensions set by $L = 10$ m while the cross-sectional area $A = 1$ m^2. The applied force F increases linearly from 0 to 0.075 N at time $t = 4$ s. The material parameters are taken as Young's modulus $E = 1000$ Pa, mass density $\rho = 1000$ kg/m^3 and internal length $\ell = 1$ m. The damage evolution of Eq. (11) is set through $\kappa_0 = 10^{-4}$ and $\kappa_u = 10^{-3}$. The time integration has been performed with a constant acceleration Newmark scheme and a time step $\Delta t = 0.2$ s. The displacements have been discretised with EFG shape functions using a quadratic base vector $\mathbf{p}$ and the ratio of domain of influence over nodal spacing $d/h = 4$. More details on the EFG discretisation can be found in [4].

In Figures 3 and 4 the results obtained with the second-order and fourth-order implicit models are shown. The evolution in time of the strain profiles along the bar, obtained with an EFG discretisation of 41 equally spaced nodes, is very similar for the two models. The damage profiles at time $t = 40$ s are compared for different discretisations (note that only the range $-10 \leq x \leq -4$ is shown). Both models converge upon refinement of the discretisation. The damage profile of the fourth-order model is slightly wider than that of the second-order model.

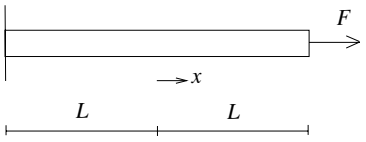


Fig. 2. One-dimensional wave propagation — problem statement.

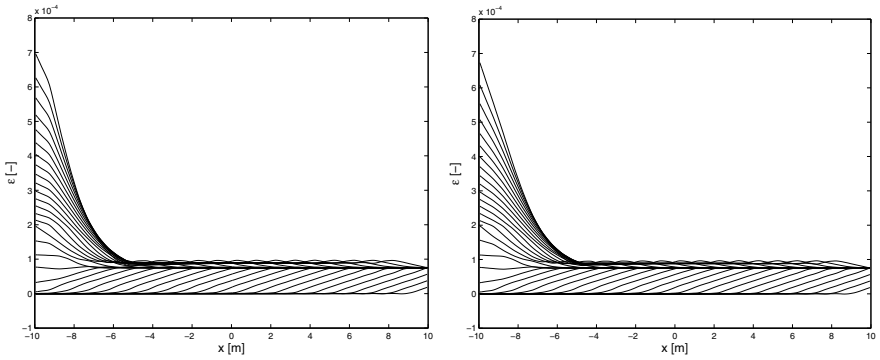


Fig. 3. One-dimensional wave propagation — stroboscopic strain evolution for second-order (left) and fourth-order (right) implicit gradient model.

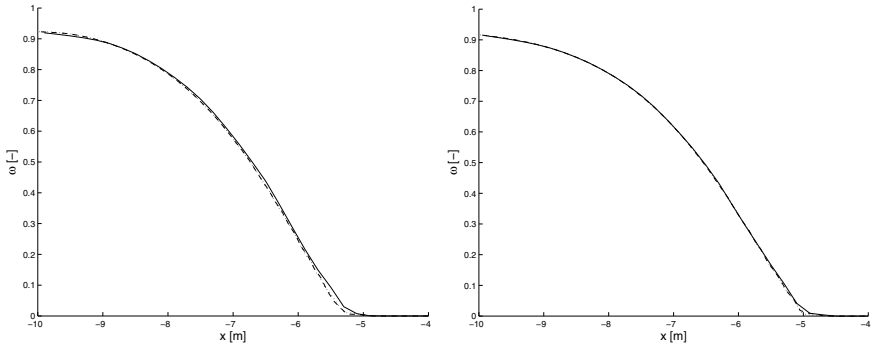


Fig. 4. One-dimensional wave propagation — convergence of final damage profiles in left part of the bar for second-order (left) and fourth-order (right) implicit gradient model; 21 nodes (solid), 41 nodes (dashed) and 81 nodes (dotted).

4.2 Notched specimen tension test

The damage evolution in a two-dimensional strip with symmetric imperfections is studied next. The test data are based on those of Chen et al. [11], see also Figure 5 where dimensions are given in m. For the bulk material the material parameters are taken as Young's modulus $E = 2 \times 10^6$ Pa and Poisson's ratio $\nu = 0$. Damage evolution is governed by $\kappa_0 = 10^{-3}$ and $\kappa_u = 0.0625$. Two symmetrically placed imperfections are modelled by reduced values for $E = 1.8 \times 10^6$ Pa, $\kappa_0 = 9 \times 10^{-4}$ and $\kappa_u = 0.056$. The bottom edge of the strip is fully fixed and the strip is loaded by imposing a vertical displacement at the top edge.

The flexible implementation offered by the EFG method is exploited by comparing the responses of gradient-type and integral-type nonlocality. Four different formulations are compared:

- the second-order implicit gradient model according to Eq. (15);

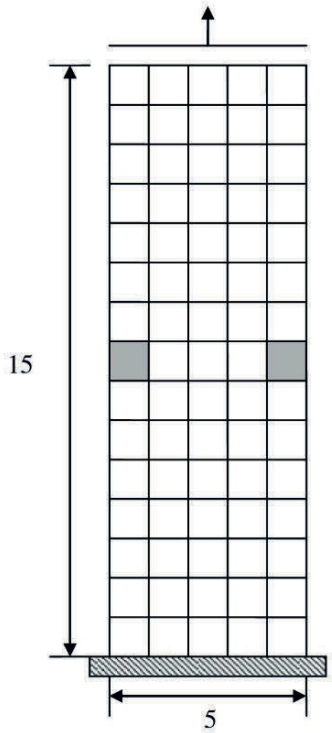


Fig. 5. Notched specimen — geometry and loading conditionsn.

- the fourth-order implicit gradient model according to Eq. (15);
- the integral nonlocal model whereby the spatial averaging operates on the equivalent strain. This is the format given in Eq. (13) and it is the original integral-type nonlocal damage model proposed in [34];
- the integral nonlocal model whereby the spatial averaging operates on the displacements. This concerns a more recent version of the integral-type damage models [21, 36, 37]. Eq. (13) is replaced by

$$\mathbf{u}^{nl}(\mathbf{x}) = \int_V \alpha(\boldsymbol{\xi})\mathbf{u}(\mathbf{x}+\boldsymbol{\xi})\,\mathrm{d}V \tag{22}$$

where $\alpha(\boldsymbol{\xi})$ is a weighting function. In order that a constant strain field results in a constant stress field the weighting function is required to have reproducibility of order 1. This implies that a linear local displacement field can be transformed into the identical nonlocal displacement field. The weighting function is accordingly constructed from the product of a first order polynomial and the normalised exponential function of Eq. (13):

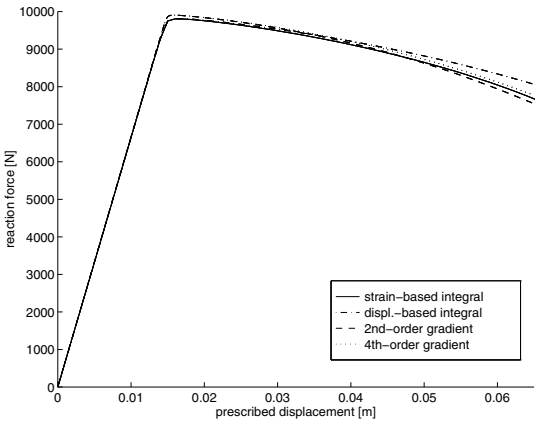


Fig. 6. Notched specimen — load-displacement curves for various nonlocal models.

$$\alpha = [c_0 + \boldsymbol{\xi}^T \mathbf{c}_1] \frac{\exp\left(-\frac{\boldsymbol{\xi}^T\boldsymbol{\xi}}{2\ell^2}\right)}{\int\limits_V \exp\left(-\frac{\boldsymbol{\xi}^T\boldsymbol{\xi}}{2\ell^2}\right) \mathrm{d}V} \tag{23}$$

The coefficients c_0 and $\mathbf{c}_1$ follow from solving a small system of linear equations as explained in detail by Rodríguez-Ferran et al. [36, 37]. After averaging the displacements, a strain field based on the nonlocal displacements is computed and used to determine the equivalent strain. The rest of the model remains unaltered.

In all cases the internal length that sets the nonlocality is taken as $\ell = 0.53$ m. The two gradient models are discretised with a 6×16 nodal distribution as shown in Figure 5, using a cubic base vector $\mathbf{p}$ and a ratio $d/h = 6$ to discretise the displacements. For the two integral models a 11×31 nodal grid with a linear base vector $\mathbf{p}$ and $d/h = 1.5$ was used. The finer nodal grid of the integral models partly balances the more extended $\mathbf{p}$-vector for the gradient models.

Figure 6 shows the load-displacement curves of the four formulations. In general an excellent agreement is found (note that this comparison would have been a tedious task if finite element implementations had been used, especially the inclusion of the fourth-order implicit gradient model). The predicted peak load is virtually the same for all formulations. Some minor differences in ductility occur whilst progressing further down the post-peak branch, but all models show an increasing downward slope which, together with a linear softening local stress-strain relation, indicates that the width of the *damaging zone* decreases with ongoing damage. Indeed, a one-dimensional dispersion analysis of the strain-based models demonstrates that the width of the damaging zone vanishes for $\omega \rightarrow 1$, which implies a smooth transition towards a discrete crack.

5 Regularising properties of EFG shape functions

Nonlocality, or interaction at finite distance, can be understood mechanically or mathematically. The former meaning is the one used in the aforementioned nonlocal gradient-type and integral-type damage models. The latter interpretation applies when discretising the governing differential equations, i.e. through replacing the spatial derivatives by a series of nodal values whereby the nodes are positioned at finite distances from one another. Both kinds of nonlocality bear similarities, the more so in case of the (mechanical) integral-type nonlocality of Eq. (13) and EFG weight functions of Eq. (8).

Chen et al. have discussed the intrinsic nonlocality of meshless methods [11] (although they refer to nonlocality as a finite distance interaction *extending beyond the nearest neighbour nodes*, not as finite distance interaction *per se*). They applied a spatial averaging in the spirit of Eq. (8) to certain state variables (namely displacements and strains). While these approaches fit well within the framework of integral-type nonlocality, the chosen terminology of "intrinsic nonlocality of meshless methods" could indicate that there is an intrinsic *mechanical* nonlocality present in meshless methods. The purpose of this section is to show that this is not the case and that mechanical nonlocality must not be confused with mathematical nonlocality.

5.1 Tapered bar in tension

A tensile bar of length 10 m is studied for which the cross-sectional area varies linearly from 10 m^2 at either end towards 9 m^2 in the centre. The Young's modulus $E = 20000$ Pa and damage growth is set by $\kappa_0 = 10^{-4}$ and $\kappa_u = 0.015$. The bar is loaded by prescribing the displacement at the right end. The EFG shape functions for the displacements are constructed with a quadratic base vector $\mathbf{p}$. The presence or absence of regularisation is tested by comparing the response of different nodal distributions.

Firstly, mechanical nonlocality is eliminated by selecting $\ell = 0$ m and instead a fixed parameter $d = 4$ m for the EFG weight function of Eq. (8) is taken. The load-displacement curves for 11, 21 and 31 equally spaced nodes are plotted in Figure 7 (left). It is clearly seen that no convergence upon discretisation refinement is obtained. Instead, every increase of the number of nodes leads to a more brittle response. Hence, it must be concluded that fixing the domain of influence radius d of the EFG shape functions is not sufficient to obtain a proper regularisation of the post-peak response.

Next, mechanical nonlocality is activated by setting $\ell = 0.35$ and the EFG radius of influence d is set as a fixed multitude of the nodal spacing h, namely $d/h = 4$. Figure 7 (right) shows the load-displacement curves for 21, 41, 81 and 161 equally spaced nodes. Although the 21 nodes response diverges somewhat from the other ones in the later loading stages, the agreement between the other three curves is excellent. This shows yet again that a proper regularisation can be obtained with mechanical nonlocality.

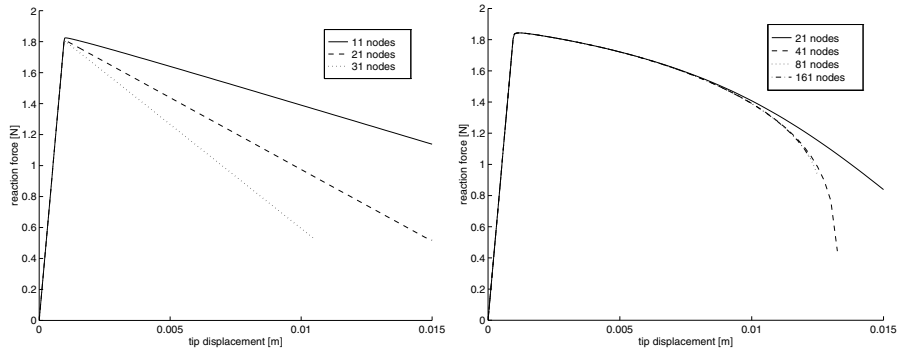


Fig. 7. Tapered bar in tension — load-displacement curves for formulations without (left) and with (right) mechanical nonlocality.

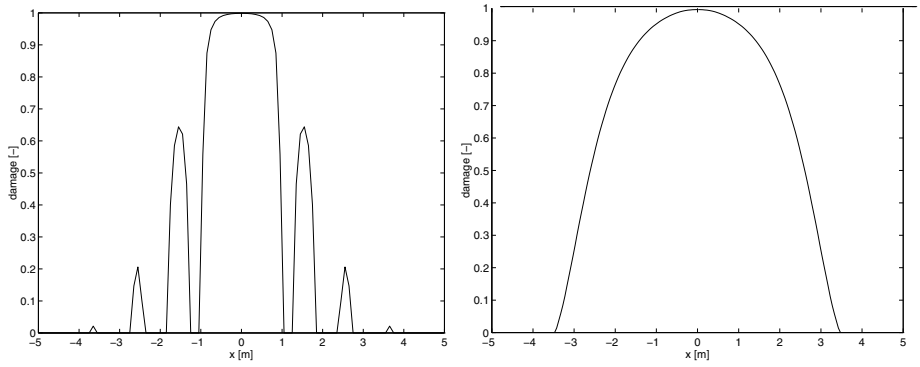


Fig. 8. Tapered bar in tension — damage profiles for 0.0125 m prescribed displacement; local model with 21 nodes (left) and nonlocal model with 41 nodes (right).

A particular observation is the differences in post-peak slope between the (mechanically) local and nonlocal models. The local models follow the slope of the local stress-strain relation, i.e. linear softening. In contrast, the nonlocal models show an increased steepness in the post-peak regime. This can be explained by the width of the zone in which damage grows. In a local model, this width is fixed by the applied discretisation and it does not change during the loading process. In contrast, the width of the damaging zone usually becomes more narrow upon further loading in a nonlocal model. In Figure 8 typical damage profiles for a local and a nonlocal model are plotted. In the local model the damaged zone has a virtually constant width irrespective of the magnitude of the damage, whereas in the nonlocal model this width diminishes gradually. Note that the local model exhibits some secondary damage zones. The strain profile of the local model tends to become discontinuous which cannot be captured by the highly smooth EFG shape functions. This causes oscillations that result in spurious secondary damage zones.

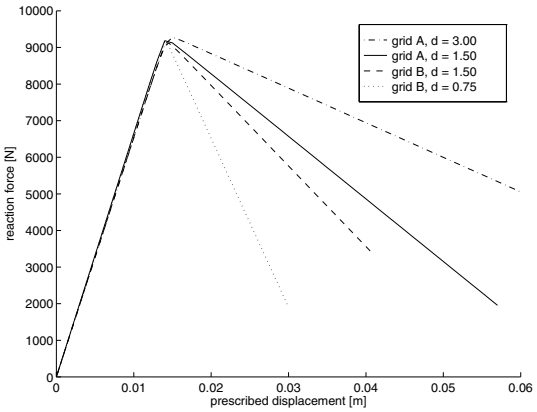


Fig. 9. Notched specimen — load-displacement curves for various local models.

5.2 Notched specimen tension test (revisited)

Finally, the test of section 4.2 is revisited without the inclusion of mechanical nonlocality, that is $\ell = 0$ m. Two different nodal distributions are considered, consisting of 6×16 nodes (grid A) and 11×31 nodes (grid B). A linear base vector $\mathbf{p}$ is used for the EFG shape functions as well as domains of influence set by $d/h = 1.5$ and $d/h = 3$ for both grids. In Figure 9 the corresponding load-displacement curves are depicted. The following can be observed:

- No convergence upon grid refinement is obtained by keeping d/h constant. For instance, compare grid A / $d = 3$ m with grid B / $d = 1.5$ m, or compare grid A / $d = 1.5$ m with grid B / $d = 0.75$ m.
- No convergence upon grid refinement is obtained by keeping d constant. For instance, compare grid A / $d = 1.5$ m with grid B / $d = 1.5$ m.

Whereas the former observation is in line with earlier observations of finite element simulations with local models [28, 38], the latter observation adds the conclusion that there is no intrinsic mechanical nonlocality in meshless methods.

6 Conclusions

In this chapter nonlocal damage models have been used in order to simulate the material nonlinearities of fracture processes in the pre-peak as well as post-peak stage. Damage models are a representation of dispersed micro-cracking and they often provide a smooth transition towards a discrete crack at the final stages of loading. The nonlocality is a key ingredient that ensures that numerical simulations converge upon refimenent of the discretisation.

Two main types of nonlocality exist: either spatial derivatives or spatial averages of the relevant state variables are taken, referred to as gradient-type and integral-type nonlocality, respectively. Both types have been proven successful in many occasions but systematic comparisons between the two types of nonlocality in a numerical context are largely lacking. Both types of models pose their own requirements on the numerical implementation, be it continuity requirements in case of gradient nonlocal models or the construction of neighbour graphs in case of integral nonlocal models. Meshless methods such as the element-free Galerkin (EFG) offer a suitable numerical platform to implement a range of nonlocal models and thus allow for in-depth analysis and comparisons.

Two sets of tests have been carried out in this chapter. In the first set, various gradient nonlocal models and integral nonlocal models have been compared. Relevant conclusions include the following:

- For the so-called implicit gradient series, there is no qualitative difference and only a marginal quantitative difference between a second-order and a fourth-order truncation. This has already been observed in statics earlier [4] and is now confirmed for a dynamic loading case.
- There is no qualitative difference and only a marginal quantitative difference between integral nonlocality operating on the equivalent strain or operating on the displacements. This confirms the findings of [36, 37].
- The parameters of the gradient-type and integral-type models can be related to each other. With such a parameter selection, the agreement between gradient-type and integral-type nonlocal models is excellent.

A second set of tests concerns the issue of absence or presence of intrinsic regularisation properties in the EFG method. Although the EFG shape functions include a certain long-range interaction that extends beyond the usual nearest neighbour interaction of finite element shape functions, it is misleading to refer to this as a mechanical nonlocality. In a series of examples it has been shown that no convergence upon refinement of discretisation is obtained if mechanical nonlocality is left out of the formulation. In particular, it has been demonstrated in a series of numerical examples that the domain of influence radius of EFG shape functions cannot be used as an intrinsic length scale. It is, however, possible to use the spatial averaging operator of meshless shape functions in the smoothing of the relevant state variables, as long as the domain of influence parameter is clearly distinguished from the internal length — this is explored in [11, 21, 36, 37].

Although the EFG method does not provide sufficient regularisation in itself, it is an excellent tool to test, analyse and compare various formats of nonlocality that do regularise damage simulations.

References

1. E.C. Aifantis. ASME J. Engng Mat. Techn., 106:326–330, 1984.
2. E.C. Aifantis. Int. J. Plast, 3:211–247, 1987.
3. H. Askes and E.C. Aifantis. Int. J. Fract., 117:347–358, 2002.
4. H. Askes, J. Pamin, and R. de Borst. Int. J. Num. Meth. Engng, 49:811–832, 2000.
5. H. Askes and L.J. Sluys. Eur. J. Mech. A/Sol., 21:379–390, 2002.
6. H. Askes and L.J. Sluys. Arch. Appl. Mech., 73:448–465, 2003.
7. Z.P. Bažant and G. Pijaudier-Cabot. ASME J. Appl. Mech., 55:287–293, 1988.
8. T. Belytschko, Y. Krongauz, D. Organ, M. Fleming, and P. Krysl. Comp. Meth. Appl. Mech. Engng, 139:3–47, 1996.
9. T. Belytschko, Y.Y. Lu, and L. Gu. Int. J. Num. Meth. Engng, 37:229–256, 1994.
10. G. Borino, B. Failla, and F. Parrinello. Int. J. Sol. Struct., 40:3621–3645, 2003.
11. J.-S. Chen, C.-T. Wu, and T. Belytschko. Int. J. Numer. Meth. Engng, 47:1303–1322, 2000.
12. C. Comi. Mech. Cohes.-Frict. Mater., 4:17–36, 1999.
13. R. de Borst, A. Benallal, and O.M. Heeres. J. Phys. IV, 6:491–502, 1996.
14. R. de Borst and H.-B. Mühlhaus. Int. J. Num. Meth. Engng, 35:521–539, 1992.
15. J.H.P. de Vree, W.A.M. Brekelmans, and M.A.J. van Gils. Comp. Struct., 55:581–588, 1995.
16. A.C. Eringen. Int. J. Engng Sci., 19:1461–1474, 1981.
17. A.C. Eringen. Int. J. Engng Sci., 21:741–751, 1983.
18. M. Frémond and B. Nedjar. Int. J. Sol. Struct., 33:1083–1103, 1996.
19. A. Huerta and G. Pijaudier-Cabot. ASCE J. Engng Mech., 120:1198–1218, 1994.
20. M. Jirásek. Int. J. Solids Struct., 35:4133–4145, 1998.
21. M. Jirásek and S. Marfia. Int. J. Num. Meth. Engng, 63:77–102, 2005.
22. M. Jirásek and B. Patzák. Comp. Struct., 80:1279–1293, 2002.
23. M. Jirásek and S. Rolshoven. Int. J. Engng Sci., 41:1553–1602, 2003.
24. T. Liebe, P. Steinmann, and A. Benallal. Comp. Meth. Appl. Mech. Engng., 190:6555–6576, 2001.
25. Y.Y. Lu, T. Belytschko, and L. Gu. Comp. Meth. Appl. Mech. Engng, 113:397–414, 1994.
26. J. Mazars and G. Pijaudier-Cabot. ASCE J. Engng Mech., 115:345–365, 1989.
27. H.-B. Mühlhaus and E.C. Aifantis. Int. J. Sol. Struct., 28:845–857, 1991.
28. J. Pamin. Gradient-dependent plasticity in numerical simulation of localization phenomena. Dissertation, Delft University of Technology, 1994.
29. J. Pamin, H. Askes, and R. de Borst. Comp. Meth. Appl. Mech. Engng, 192:2377–2403, 2003.
30. R.H.J. Peerlings, R. de Borst, W.A.M. Brekelmans, and J.H.P. de Vree. Int. J. Num. Meth. Engng, 39:3391–3403, 1996.
31. R.H.J. Peerlings, R. de Borst, W.A.M. Brekelmans, J.H.P. de Vree, and I. Spee. Eur. J. Mech. A/Sol., 15:937–953, 1996.
32. R.H.J. Peerlings, M.G.D. Geers, R. de Borst, and W.A.M. Brekelmans. Int. J. Sol. Struct., 38:7723–7746, 2001.
33. R.H.J. Peerlings, T.J. Massart, and M.G.D. Geers. Comp. Meth. Appl. Mech. Engng, 193:3403–3417, 2004.
34. G. Pijaudier-Cabot and Z.P. Bažant. ASCE J. Engng Mech., 113:1512–1533, 1987.

35. G. Pijaudier-Cabot and A. Huerta. Comp. Meth. Appl. Mech. Engng, 90:905–919, 1991.
36. A. Rodríguez-Ferran, I. Morata, and A. Huerta. Comp. Meth. Appl. Mech. Engng, 193:3431–3455, 2004.
37. A. Rodríguez-Ferran, I. Morata, and A. Huerta. Int. J. Num. An. Meth. Geom., 29:473–493, 2005.
38. L.J. Sluys. Wave propagation, localisation and dispersion in softening solids. Dissertation, Delft University of Technology, 1992.
39. Z. Tang, S. Shen, and S.N. Atluri. Comp. Model. Engng Sci., 4:177–196, 2003.

Three-dimensional non-linear fracture mechanics by enriched meshfree methods without asymptotic enrichment

Stéphane Bordas[1], Goangseup Zi[2], and Timon Rabczuk[3]

[1] Assistant professor (lecturer), University of Glasgow, Civil Engineering department, Rankine building, G12 8LT, Glasgow, UK bordas@civil.gla.ac.uk http://www.civil.gla.ac.uk/~bordasTel: +44(0)141 330 4075

[2] Assistant Professor, Department of Civil & Environmental Engineering, Korea University, 5 Ga 1, An-Am Dong, Sung-Buk Gu, Seoul, 136-701, Korea; Tel.: +82-2-3290-3324; g-zi@korea.ac.kr http://www.korea.ac.kr/~semlab/.

[3] Senior Lecturer, Dep. Mechanical Engineering, University of Canterbury, Christchurch, New Zealand; timon.rabczuk@canterbury.ac.nz http://www.lnm.mw.tum.de/Members/rabczuk

Summary. This paper presents a three-dimensional, extrinsically enriched meshfree method for initiation, growth and coalescence of an arbitrary number of cracks in non-linear solids including large deformations, for statics and dynamics.

The novelty of the methodology fashioned in this work is that only an extrinsic discontinuous enrichment and no near-tip/asymptotic enrichment is required. Instead, a Lagrange multiplier field is added along the crack front to close the crack along the front. This decreases the computational cost and removes difficulties involved with a branch enrichment.

Numerical examples treated include the pull-out of a reinforcement bar from a concrete block, and a Taylor bar impact with very large deformation and fragmentation. The results are compared to experimental results, and other simulations from the literature, which shows the robustness and accuracy of the method.

Key words: eXtended Element Free Galerkin Method (XEFG); discontinuous enrichment; Cohesive cracks; Lagrange multipliers; Dynamic fracture; large deformations; high velocity impact

1 Introduction

We present a simple enriched meshfree method to model three-dimensional crack initiation, propagation and coallescence, including branching in non-linear materials for statics and dynamics. The major advantage of the technique is that no near-front enrichment is required.

Alain Combescure et al. (eds.), IUTAM Symposium on Discretization Methods for Evolving Discontinuities, 21–36.

In finite element methods, either cracks lie along element boundaries, as in Xu and Needleman (1994), Camacho and Ortiz (1996), Ortiz and Pandolfi (1999), Zhou and Molinari (2004) or the discontinuities are introduced through the elements. Two classes of techniques have been developed to introduce this discontinuity inside finite elements: the strong discontinuity approach such as the extended finite element method (XFEM) Belytschko and Black (1999), Moës et al. (1999) and the weak discontinuity approach. When cracks are not aligned with element boundaries, their geometry must be followed independently of the mesh. The level set method is a possible technique to track the geometry of the cracks, but for multiple cracks, it is not well adapted. Alternatively, the cracks can be handled geometrically Duflot (2006). In 3D, this requires updating triangulations of the cracks as they evolve.

The method we are using is based on the Element-Free Galerkin method (EFG) introduced by Belytschko et al. (1994). For such methods, no discretization mesh is present, but a background mesh used for integration purposes is usually present. The cracks can be represented by the level set or alternate vector level set methods Ventura et al. (2002, 2003). It can alternatively be discretized (triangulated) Duflot (2006). Yet another approach, used here, is to define the cracks through the background integration mesh Rabczuk et al. (2007).

Among the strong discontinuity approaches, the extended finite element method (XFEM) is one of the most versatile and accurate. This method has been successfully applied to static problems in two and three dimensions, (see e.g Moës et al. (1999, 2002), Zi and Belytschko (2003), Zi et al. (2004), Gravouil et al. (2002), Moës et al. (2002)) and to dynamic problems Belytschko et al. (2003), Zi et al. (2005) in two dimensions and three dimensions including contact along the crack faces, and small-scale plasticity Rethore et al. (2005). The extended finite element method has also been utilised in industrial settings to assess damage tolerance of complex structures by Bordas and Moran (2006), Bordas et al. (2006a) and open source C++ libraries are available, such as two libraries developed by Bordas et al. (2006b) and Dunant et al. (2006).

An advantage of meshfree methods is their higher continuity, smoothing the stress field in the crack tip region. Work such as Belytschko and Tabbara (1996), Lu et al. (1995), Belytschko and Lu (1995), Belytschko et al. (1995) and Krysl and Belytschko (1999) handle two and three dimensional fracture mechanics, and treat discontinuities by the visibility criterion or some modifications of it. The support of the shape functions is truncated by the line of discontinuity. Other novel approaches can treat kinked and curved discontinuities Ventura et al. (2002), where the moving least squares (MLS) basis functions around the crack tip is shown to improve convergence. The major drawback is the need for an explicit representation of the crack. Alternate methods enrich the MLS weight functions Duflot (2006) and use the diffraction criterion to introduce the discontinuity.

Rabczuk et al. (2007) propose an extended element free Galerkin (XEFG) method for cohesive crack initiation, growth and junction in three dimensional statics and dynamics, but the closure of the crack along the front is ensured through near-tip enrichment vanishing along the front. The polar coordinate system along a crack front is not well-defined at the kinks along this front, as made clear in Rabczuk et al. (2007). Another difficulty with near-front enrichment is that the "correct" fields are neither known for large strain nor for non-linear materials in general. This makes the selection of these fields difficult and calls for an alternate method.

In the method we propose, front enrichment is unnecessary, and only discontinuous enrichment is used; the crack is closed using a constraint field enforced by Lagrange multipliers. Similar accuracy compared to Rabczuk et al. (2007) is obtained on the examples treated.

The outline of the paper is as follows: we describe the approximation of the jump in displacement in the meshfree method employed, method which we then briefly recall. We state the governing equations, their weak and discretized form, before describing the crack modelling algorithm. The method is then exercised on a quasi-static and dynamic crack growth problem.

2 Element Free Galerkin (EFG) Approximation

The construction of the moving least squares (MLS) approximation used to describe the displacement field in the element-free Galerkin method is now standard and the interested readed is referred to Belytschko and Tabbara (1996), Belytschko et al. (1994), Fleming et al. (1997). In the whole paper, we denote by Ω_0 the body under interest in the reference configuration, and $\mathbf{X} = (X, Y, Z) \in \Omega_0 \subset \mathbb{R}^3$ a point in the reference configuration. The displacement field is approximated by moving least squares. We use the cubic B-spline, circular supports and a linear basis $\mathbf{p} = [1, X, Y, Z]$. Denote by $\mathcal{I}_{I,\mathbf{X}}$ the set of $N_{I,\mathbf{X}}$ particles whose domain of influence contain a given point $\mathbf{X} \in \Omega_0$. For all the $N_I(\mathbf{X})$ particles whose domain of influence contains a given point $\mathbf{X} \in \Omega_0$, the MLS shape function Φ_I associated with particle $I = \mathbf{X}_I$ evaluated at point $\mathbf{X}$ is obtained by the matrix/vector product

$$\underbrace{\Phi_I(\mathbf{X})}_{N_{I,\mathbf{X}}\times 1} = \underbrace{\mathbf{p}^T}_{N_{I,\mathbf{X}}\times 4}(\mathbf{X}) \cdot \underbrace{\mathbf{A}(\mathbf{X},\mathbf{X}_I)^{-1}}_{4\times 4} \cdot \underbrace{\mathbf{D}(\mathbf{X},\mathbf{X}_I)}_{4\times 1} \tag{1}$$

$$\mathbf{A}(\mathbf{X},\mathbf{X}_I) = \sum_{I\in\mathcal{I}_{I,\mathbf{X}}} \mathbf{p}(\mathbf{X}_I)\ \mathbf{p}^T(\mathbf{X}_I)\, W(\bar{r}_I(\mathbf{X})) \tag{2}$$

$$\mathbf{D}(\mathbf{X},\mathbf{X}_I) = \sum_{I\in\mathcal{I}_{I,\mathbf{X}}} \mathbf{p}(\mathbf{X}_I)\, W(\bar{r}_I(\mathbf{X})) \tag{3}$$

With these definitions, the standard part (C_2-continuous and linear-complete) of the MLS approximation (denoted by a superscript ^{s}) for the displacement field $\mathbf{u}^s$ at time t can be written as

$$\forall \mathbf{X} \in \Omega_0, \forall t > 0, \quad \mathbf{u}^s(\mathbf{X}, t) = \sum_{I \in \mathcal{W}} \Phi_I(\mathbf{X})\ \mathbf{u}_I^s(t) \tag{4}$$

3 Discontinuous enrichment of the displacement Field

We write the displacement approximation is written as the sum of a standard ($\mathbf{u}^s$) part as in Eq. (4) and an enriched part ($\mathbf{u}^e$), discontinuous through the crack faces

$$\forall \mathbf{X} \in \Omega_0, \forall t > 0, \quad \mathbf{u}(\mathbf{X}, t) = \mathbf{u}^s(\mathbf{X}, t) + \mathbf{u}^e(\mathbf{X}, t) \tag{5}$$

where $\mathbf{u}^s$ is the continuous displacement field defined by (4) and $\mathbf{u}^e$ is the discontinuous (or the enriched) displacement field defined below. Define $\mathcal{E}$ as the set of all the cracks in the domain and $\mathbf{u}_\alpha^e$ the enriched part of the approximation due to crack α. The enriched part of the approximation due to all cracks in $\mathcal{E}$ writes

$$\forall \mathbf{X} \in \Omega_0, \forall t > 0, \quad \mathbf{u}^e(\mathbf{X}, t) = \sum_{\alpha \in \mathcal{E}} \mathbf{u}_\alpha^e(\mathbf{X}, t) \tag{6}$$

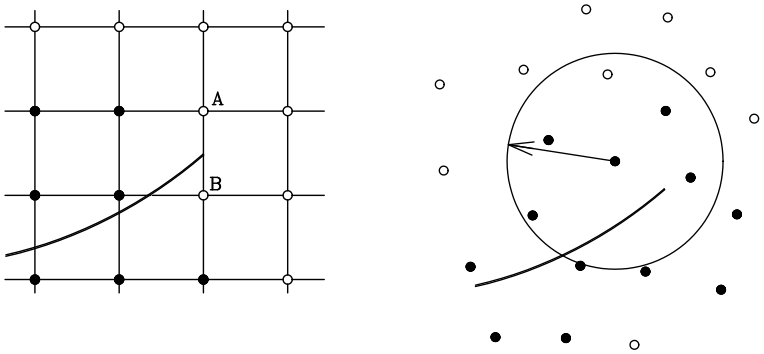
Let $\mathcal{W}_\alpha$ be the set of particles whose domain of influence is cut by crack α, and Ψ_I^α be the enrichment function associated with particle I and crack α, discontinuous through this crack and defined below. Define by $\boldsymbol{a}_I^\alpha$ the additional degrees of freedom for the enrichment Ψ_I^α. The discontinuous part of the displacement approximation due to crack $\alpha \in \mathcal{E}$ writes

$$\forall \mathbf{X} \in \Omega_0, \forall t > 0, \quad \mathbf{u}_\alpha^e(\mathbf{X}, t) = \sum_{I \in \mathcal{W}_\alpha} \Phi_I(\mathbf{X})\ \Psi_I^\alpha(\mathbf{X})\ \boldsymbol{a}_I^\alpha(t) \tag{7}$$

We now define the enrichment functions Ψ_I^α. If a domain of influence is cut by a crack, it is enriched with the sign function given by

$$\text{sign}(x) = \begin{cases} 1 \ \text{ for } x > 0 \\ \\ -1 \text{ for } x < 0 \end{cases} \tag{8}$$

Let $\mathbf{n}$ be the crack normal and Γ_c^α represent the surface of crack α. Let $\mathbf{X}_\alpha$ be a point on the surface Γ_c^α of crack $\alpha \in \mathcal{E}$ and $\mathbf{X}$ a point in Ω_0. The sign of quantity $\mathbf{n} \cdot (\mathbf{X} - \mathbf{X}_\alpha)$ defines on which side of Γ_c^α $\mathbf{X}$ is located. The distance from $\mathbf{X}$ to Γ_c^α is $\min_{\mathbf{X}^\Gamma \in \Gamma_c^\alpha} \|\mathbf{X} - \mathbf{X}^\Gamma\|$. The signed distance $f^\alpha(\mathbf{X})$ from $\mathbf{X}$ to crack α is



(a) finite element method (b) meshless method

Fig. 1. The enrichment for the crack tip by using the step function in (a) the finite element method and (b) meshless methods; solids are enriched nodes and circles unenriched nodes.

$$\forall \mathbf{X} \in \Omega_0, \forall \mathbf{X}_\alpha \in \Gamma_c^\alpha, \quad f^\alpha(\mathbf{X}) = \text{sign}\left[\mathbf{n} \cdot (\mathbf{X} - \mathbf{X}_\alpha)\right] \min_{\mathbf{X}^\Gamma \in \Gamma_c^\alpha} \|\mathbf{X} - \mathbf{X}^\Gamma\| \quad (9)$$

The enrichment for crack α and particle I writes

$$\forall \alpha \in \mathcal{E}, \forall I \in \mathcal{W}, \forall \mathbf{X} \in \Omega_0, \quad \Psi_I^\alpha(\mathbf{X}) = \text{sign}\left[f^\alpha(\mathbf{X})\right] - \text{sign}\left[f^\alpha(\mathbf{X}_I)\right] \quad (10)$$

3.1 Closing the crack along the front

Crack closure along the front is a natural outcome of near-front enrichment vanishing along the crack front, as in (Rabczuk et al., 2007). In the context of the extended finite element method, a methodology was devised to close the crack without near-front enrichment (Dolbow and Devan, 2003, Zi and Belytschko, 2003) as long as the crack tip is located on element edges. This idea does not apply to meshfree methods because of the strong overlapping of domains of influence. In C_0-FEM-based methods, the shape function associated with each node is only coupled with those of the nodes contained in its support. Imagine a crack tip in a two-dimensional finite element mesh; see Figure 1a, the crack tip is positioned on the edge connecting nodes A and B. Because the crack must close along the front, nodes A and B should not be enriched. Figure 1b shows the case of meshless methods. The domain of influence for a particle in meshless methods overlaps heavily with that of other particles.

We propose the construction of a Lagrange multiplier field to close the cracks along their fronts. The idea is an extension of the principle presented

in Carpinteri et al. (2001). If only the sign function enrichment of Eq. (10) is used, the discontinuity Γ_c extends beyond the crack front. To correctly model the crack, the discontinuity on the front Γ_c should vanish. Because the condition should be satisfied along a surface, the Lagrange multiplier field must be discretized and we use the same shape functions as those for the domain partially cut by the crack.

4 Description of Cracks

4.1 Geometric description

The cracks are described similarly as in Rabczuk et al. (2007). The crack surfaces are non-planar and represented by the union of planar segments obtained by slicing the tetrahedral background mesh by the failure planes obtained through material stability analysis (described in Section 4.2). The representation of crack junction through signed-distance functions is now standard, and the interested reader is referred to Rabczuk et al. (2007) for details.

4.2 Initiation and propagation of cracks

We employ the loss of hyperbolicity (in dynamics) and loss of ellipticity (in quasi-statics) criterion for crack initiation and propagation, as proposed by Belytschko et al. (2003), Rabczuk and Zi (2006), Rabczuk et al. (2007). Therefore, a crack is initiated or propagated if the minimum eigenvalue of the acoustic tensor $\boldsymbol{Q}$ is smaller or equal to zero:

$$\min \text{eig}(\boldsymbol{Q}) \leq 0 \text{ with } \quad \boldsymbol{Q} = \mathbf{n} \cdot \mathbf{A} \cdot \mathbf{n} \tag{11}$$

where $\mathbf{n} = (cos\alpha\ cos\varphi, cos\alpha\ sin\varphi, sin\alpha)$ is the normal to the crack surface depending on the angle θ, $\mathbf{A} = \boldsymbol{C}^t + \boldsymbol{\sigma} \otimes \boldsymbol{I}$, $\boldsymbol{\sigma}$ is the stress tensor and $\boldsymbol{C}^t$ is the fourth order tangential modulus.

4.3 Tracking the crack path

The algorithmic procedure to track the crack path is detailed in Rabczuk et al. (2007). We now allow multiple cracks to overlap during initiation (at the same time step[4]). Especially for problems with excessive cracking, this facilitates the implementation and also decreases computational cost.

[4] in contrast to our method in Rabczuk et al. (2007) where we did not allow overlapping of simultaneously initiated cracks

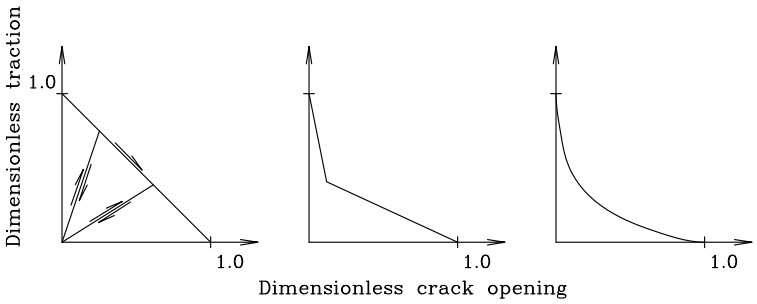


Fig. 2. The types of the cohesive laws frequently used in practice; (a) linear (or triangular), (b) bilinear and (c) exponential cohesive laws.

5 Governing Equations

5.1 The momentum equation and the boundary conditions

Defining Γ_0^c as the union of all the crack surfaces, the strong form of the momentum equation in the total Lagrangian description is given by

$$\varrho_0 \, \ddot{\mathbf{u}} = \boldsymbol{\nabla}_{s0} \cdot \mathbf{P} + \varrho_0 \, \boldsymbol{b} \quad \text{in} \quad \Omega_0 \setminus \Gamma_0^c \tag{12}$$

with boundary conditions:

$$\mathbf{u}(\mathbf{X}, t) = \bar{\mathbf{u}}(\mathbf{X}, t) \quad \text{on} \quad \Gamma_0^u \tag{13}$$

$$\mathbf{n}_0 \cdot \mathbf{P}(\mathbf{X}, t) = \bar{\boldsymbol{t}}(\mathbf{X}, t) \quad \text{on} \quad \Gamma_0^t \tag{14}$$

$$\mathbf{n}_0 \cdot \mathbf{P}^- = \mathbf{n}_0 \cdot \mathbf{P}^+ = \boldsymbol{t}_{c0} \quad \text{on} \quad \Gamma_0^c \tag{15}$$

$$\boldsymbol{t}_{c0} = \boldsymbol{t}_{c0}(\llbracket \mathbf{u} \rrbracket) \quad \text{on} \quad \Gamma_0^c \tag{16}$$

where ϱ_0 is the initial mass density, $\ddot{\mathbf{u}}$ is the acceleration, $\mathbf{P}$ denotes the nominal stress tensor, $\boldsymbol{b}$ designates the body force, $\bar{\mathbf{u}}$ and $\bar{\boldsymbol{t}}$ are the prescribed displacement and traction, respectively, $\mathbf{n}_0$ is the outward normal to the domain and $\Gamma_0^u \cup \Gamma_0^t \cup \Gamma_0^c = \Gamma_0$, $(\Gamma_0^u \cap \Gamma_0^t) \cup (\Gamma_0^t \cap \Gamma_0^c) \cup (\Gamma_0^c \cap \Gamma_0^u) = \emptyset$. Moreover, we assume that the stresses $\mathbf{P}$ are bounded on the crack surface Γ_0^c. Since the stresses are not well defined along the crack, the crack surface Γ_0^c is excluded from the domain Ω_0 which is considered as an open set.

5.2 Cohesive cracks

We exclusively used initially rigid cohesive models as shown in Figure 2 for the one-dimensional case. If a potential for the cohesive crack is defined, the unidirectional relation can be extended to general mixed mode problems, as in (Bažant and Caner, 2005, Camacho and Ortiz, 1996, Ortiz and Pandolfi,

1999). Note that if only one-dimensional cohesive models for mode-I fracture are used, traction continuity is in general violated. We only used cohesive models that fulfill the traction continuity condition, i.e. the cohesive traction $\mathbf{t}_c = \mathbf{n} \cdot \boldsymbol{\sigma}$ at crack initiation is compatible with the stress state at crack initiation. Difficulties occur at crack initiation due to the infinite slope of the tangent, especially when unloading occurs at an early stage. We note that the fulfillment of traction continuity is a complementary problem and cannot simply be enforced by Lagrange multipliers since the traction cannot exceed the tensile strength of the material.

6 Discretized Equations

The weak form of the momentum equation is given by

$$\delta W = \delta W_{\text{int}} + \delta W_{\text{kin}} - \delta W_{\text{ext}} - \delta W_{\text{coh}} + \delta W_{\text{L}} \tag{17}$$

in which $\delta W_{\text{int}}, \delta W_{\text{kin}}, \delta W_{\text{ext}}, \delta W_{\text{coh}}$ are the parts composing the virtual work of the internal stress, the inertia force, the external traction and the cohesive traction, respectively; δW_{L} is introduced to close the crack at its crack front. Noting $\boldsymbol{\nabla}_{s0}$ the symmetric gradient operator in the reference configuration, the four parts of the virtual work write[5]

$$\delta W_{\text{int}} = \int_{\Omega_0 \setminus \Gamma_0^c} (\boldsymbol{\nabla}_{s0} \otimes \delta \mathbf{u})^T : \mathbf{P} \, \mathrm{d}\Omega_0 \tag{18}$$

$$\delta W_{\text{kin}} = \int_{\Omega_0 \setminus \Gamma_0^c} \varrho_0 \, \delta \mathbf{u} \cdot \ddot{\mathbf{u}} \, \mathrm{d}\Omega_0 \tag{19}$$

$$\delta W_{\text{ext}} = \int_{\Omega_0 \setminus \Gamma_0^c} \varrho_0 \, \delta \mathbf{u} \cdot \boldsymbol{b} \, \mathrm{d}\Omega_0 + \int_{\Gamma_0^t} \delta \mathbf{u} \cdot \bar{\boldsymbol{t}}_0 \, \mathrm{d}\Gamma_0 \tag{20}$$

$$\delta W_{\text{coh}} = \int_{\Gamma_0^c} [\![\delta \mathbf{u}]\!] \cdot \boldsymbol{\tau} \, \mathrm{d}\Gamma_0 \tag{21}$$

$$\delta W_{\text{L}} = \delta(\boldsymbol{\Lambda} \cdot [\![\mathbf{u}]\!]) \tag{22}$$

in which δW_{L} is the general variation with constraint and $\boldsymbol{\Lambda}$ is the Lagrange multiplier vector. As the Lagrange multiplier is defined for $\Gamma_{\text{c,ext}}$ and discretized using the shape functions of the meshless method, the discretized Lagrange multiplier is given by

$$\boldsymbol{\Lambda} = \boldsymbol{\Phi}^0 \boldsymbol{\lambda} \tag{23}$$

where $\boldsymbol{\lambda}$ are the coefficients of the discretization of the Lagrange multiplier Λ. Substituting the continuous and discontinuous displacement fields $\mathbf{u}^s$ and

[5] recall that Γ_0^c is the union of all crack surfaces.

$\mathbf{u}^e$ in Eqs. (4) and (6), and the crack opening displacement $[\![\mathbf{u}]\!]$ to the weak form, we obtain[6]

$$\begin{aligned}
\delta W_{\text{int}} &= \sum_{I\in\mathcal{W}} \delta\mathbf{u}_I^T \int_{\Omega_0\setminus\Gamma_0^c} \boldsymbol{\nabla}_{s0}\boldsymbol{\Phi}_I(\mathbf{X})^T : \mathbf{P}\,\mathrm{d}\Omega \\
&\quad + \sum_{\alpha\in\mathcal{E}}\sum_{K\in\mathcal{W}_\alpha} \delta\boldsymbol{a}_K^{\alpha\,T} \int_{\Omega_0\setminus\Gamma_0^{c\alpha}} \boldsymbol{\nabla}_{s0}\left[\boldsymbol{\Phi}_K(\mathbf{X})\,\Psi_K^\alpha(\mathbf{X})\right]^T : \mathbf{P}\,\mathrm{d}\Omega \qquad (24)\\
\delta W_{\text{kin}} &= \sum_{I\in\mathcal{W}} \delta\mathbf{u}_I^T \sum_{J\in\mathcal{W}} \int_{\Omega_0\setminus\Gamma_0^c} \varrho_0\,\boldsymbol{\Phi}_I(\mathbf{X})^T\cdot\boldsymbol{\Phi}_J(\mathbf{X})\,\mathrm{d}\Omega\;\ddot{\mathbf{u}}_J \\
&\quad + \sum_{I\in\mathcal{W}} \delta\mathbf{u}_I^T \sum_{\alpha\in\mathcal{E}}\sum_{K\in\mathcal{W}_\alpha} \int_{\Omega_0\setminus\Gamma_0^{c\alpha}} \varrho_0\,\boldsymbol{\Phi}_I(\mathbf{X})^T\cdot\boldsymbol{\Phi}_K(\mathbf{X})\Psi_K^\alpha(\mathbf{X})\,\mathrm{d}\Omega\;\ddot{\boldsymbol{a}}^a l_K \\
&\quad + \sum_{\alpha\in\mathcal{E}}\sum_{K\in\mathcal{W}_\alpha} \delta\boldsymbol{a}_K^{\alpha\,T} \sum_{I\in\mathcal{W}} \int_{\Omega_0\setminus\Gamma_0^{c\alpha}} \varrho_0\,\left[\boldsymbol{\Phi}_K(\mathbf{X})\Psi_K^\alpha(\mathbf{X})\right]^T\cdot\boldsymbol{\Phi}_I(\mathbf{X})\,\mathrm{d}\Omega\;\ddot{\mathbf{u}}_I \\
&\quad + \sum_{\alpha\in\mathcal{E}}\sum_{K\in\mathcal{W}_\alpha} \delta\boldsymbol{a}_K^{\alpha\,T} \sum_{\beta\in\mathcal{E}}\sum_{M\in\mathcal{W}_\beta} \qquad (25)\\
&\qquad \int_{\Omega_0\setminus(\Gamma_0^{c\beta}\bigcup\Gamma_0^{c\alpha})} \varrho_0\,\left[\boldsymbol{\Phi}_K(\mathbf{X})\Psi_K^\alpha(\mathbf{X})\right]^T\cdot\boldsymbol{\Phi}_M(\mathbf{X})\Psi_M^\beta(\mathbf{X})\,\mathrm{d}\Omega\;\ddot{\boldsymbol{a}}_M^L \\
\delta W_{\text{ext}} &= \sum_{I\in\mathcal{W}} \delta\mathbf{u}_I^T \int_{\Omega_0\setminus\Gamma_0^c} \varrho_0\,\boldsymbol{\Phi}_I(\mathbf{X})^T\cdot\boldsymbol{b}\,\mathrm{d}\Omega \qquad (26)\\
&\quad + \sum_{I\in\mathcal{W}} \delta\mathbf{u}_I^T \int_{\Gamma_0^t} \boldsymbol{\Phi}_I(\mathbf{X})^T\cdot\bar{\boldsymbol{t}}_0\,\mathrm{d}\Gamma \\
\delta W_{\text{coh}} &= 2\sum_{\alpha\in\mathcal{E}}\sum_{K\in\mathcal{W}_\alpha} \delta\boldsymbol{a}_K^T \int_{\Gamma_0^{c\alpha}} \boldsymbol{\Phi}_K(\mathbf{X})^T\cdot\bar{\boldsymbol{t}}_c\;\mathrm{d}\Gamma \qquad (27)
\end{aligned}$$

Using the fundamental lemma of calculus, we obtain the discretized equations

$$\mathbf{M}\;\ddot{\mathbf{q}} = \mathbf{f}_{\text{ext}} + \mathbf{f}_{\text{coh}} - \mathbf{f}_{\text{int}} - \mathbf{f}_{\text{con}} \qquad (28)$$

$$\mathbf{G}\;\mathbf{a} = \mathbf{0} \qquad (29)$$

where $\mathbf{M}$ is the consistent mass matrix, $\mathbf{q}$ is the generalized parameters, $\mathbf{f}_{\text{ext}}, \mathbf{f}_{\text{int}}, \mathbf{f}_{\text{coh}}$ are the discrete external, internal and cohesive force vectors, respectively; $\mathbf{f}_{\text{con}} = \boldsymbol{\lambda}^T\mathbf{G}$ in which $\mathbf{G} = 2\;\boldsymbol{\Phi}^T\boldsymbol{\Phi}^e$ is the force term due to the constraint to close the crack at its front. The expressions for $\mathbf{M}$, $\mathbf{q}$, $\mathbf{f}_{\text{ext}}$, $\mathbf{f}_{\text{coh}}$ and $\mathbf{f}_{\text{coh}}$ are given by

[6] we denote by $\Gamma_0^{c\alpha}$ the surface of crack α in the reference configuration. The union of the $\Gamma_0^{c\alpha}$ for $\alpha\in\mathcal{E}$ forms set Γ_0^c.

$$\mathbf{M} = \int_{\Omega_0 \setminus \Gamma_0^c} \varrho_0 \begin{bmatrix} \mathbf{\Phi}^T \mathbf{\Phi} & \mathbf{\Phi}^T \mathbf{\Phi}^e \\ \mathbf{\Phi}^{e\,T} \mathbf{\Phi} & \mathbf{\Phi}^{e\,T} \mathbf{\Phi} \end{bmatrix} \mathrm{d}\Omega \tag{30}$$

$$\mathbf{f}_{\text{int}} = \int_{\Omega_0 \setminus \Gamma_0^c} \mathbf{B}^T \mathbf{P}\, \mathrm{d}\Omega + \int_{\Omega_0 \setminus \Gamma_0^c} \mathbf{B}^{e\,T} \mathbf{P}\, \mathrm{d}\Omega \tag{31}$$

$$\mathbf{f}_{\text{ext}} = \int_{\Omega_0 \setminus \Gamma_0^c} \varrho_0\, \mathbf{\Phi}^T \mathbf{b}\, \mathrm{d}\Omega + \int_{\Gamma_0^t} \mathbf{\Phi}^T \bar{\mathbf{t}}_0\, \mathrm{d}\Gamma \tag{32}$$

$$\mathbf{f}_{\text{coh}} = 2 \int_{\Gamma_0^c} \mathbf{\Phi}^T \mathbf{t}_{c0}\, \mathrm{d}\Gamma \tag{33}$$

$$\mathbf{q} = \begin{Bmatrix} \mathbf{u} \\ \mathbf{a} \end{Bmatrix} \tag{34}$$

$$\mathbf{u} = [\,\mathbf{u}_I^s\,]^T \quad \forall\ I \in \mathcal{W} \ \text{ and } \ \mathbf{a} = [\boldsymbol{a}_K^\alpha]^T \quad \forall\ K \in \mathcal{W}_\alpha,\ \forall\ \alpha \in \mathcal{E} \tag{35}$$

$$\mathbf{\Phi} = [\boldsymbol{\Phi}_I] \quad \forall\ I \in \mathcal{W} \ \text{ and } \ \mathbf{\Phi}^e = [\boldsymbol{\Phi}_\alpha \boldsymbol{\Psi}_K^\alpha] \quad \forall\ K \in \mathcal{W}_\alpha,\ \forall\ \alpha \in \mathcal{E} \tag{36}$$

$$\mathbf{B} = \boldsymbol{\nabla}_{s0} \mathbf{\Phi} \ \text{ and } \ \mathbf{B}^e = \boldsymbol{\nabla}_{s0} \mathbf{\Phi}^e \tag{37}$$

7 Numerical Examples

7.1 Pull-out test

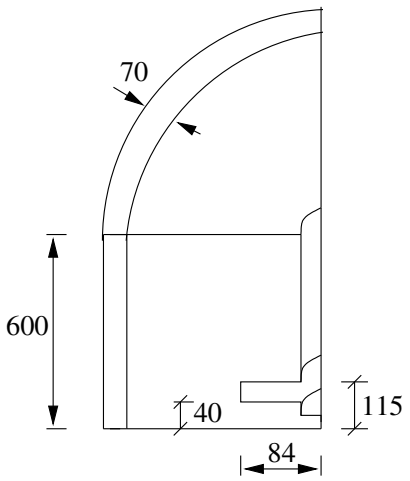


Fig. 3. Test-setup of the Pull out test

Consider a pull out test of reinforced concrete as shown in Figure 3. This example was studied previously by Gasser and Holzapfel (2005) and Areias and Belytschko (2005) by the PUFEM and XFEM, respectively. We also employed symmetry conditions and modelled only one quarter of the specimen. A vertical displacement boundary condition is applied to pull the reinforcement bar out of the concrete specimen as illustrated in Figure 3. We adopted the same constitutive and cohesive model as Gasser and Holzapfel (2005). The material parameters are $\kappa = 16,670\text{MPa}$ and $\nu = 12,500\text{MPa}$. For the cohesive model, we use, according to Gasser and Holzapfel (2005), $t_0 = 3\text{MPa}$, $a = 11.323\text{mm}^{-1}$, $b = 0.674$ and $\alpha = 1$.

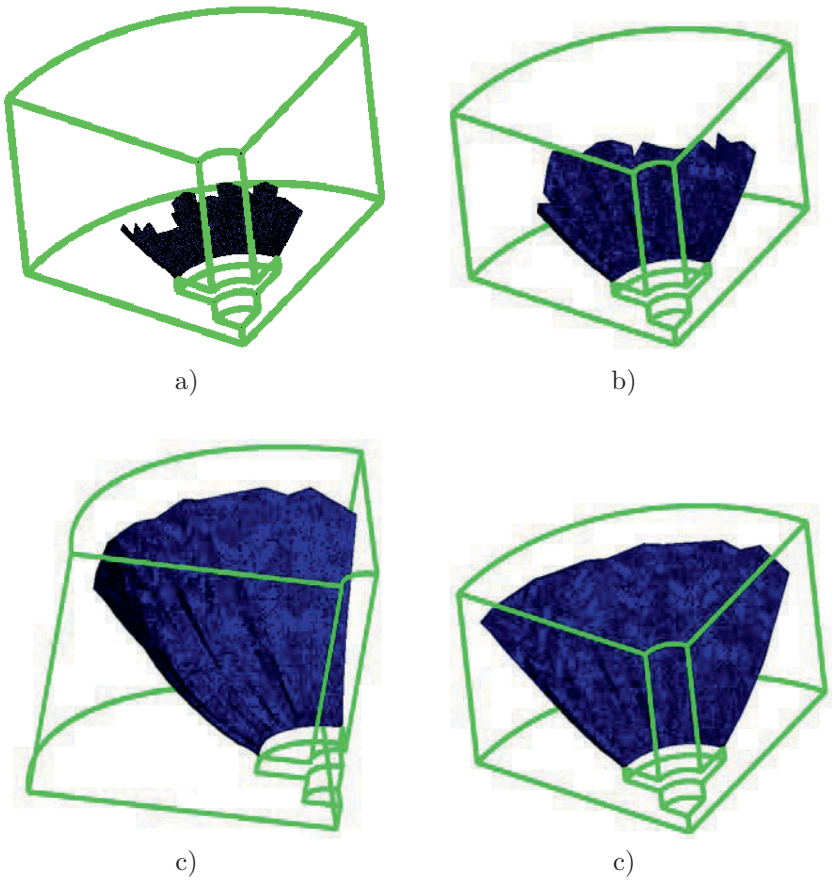


Fig. 4. Crack pattern of the pull-out test

We tested two discretizations starting with 15,000 and 45,000 nodes and refined adaptively Rabczuk and Belytschko (2005). The crack pattern at different load stages is shown in Figure 4 for the fine discretization and different view points. Note that the surface of the crack is rippling, and is captured nicely. The load-deflection curve is shown in Figure 5 and is similar to results presented in Gasser and Holzapfel (2005) and Areias and Belytschko (2005) as well as Rabczuk et al. (2007).

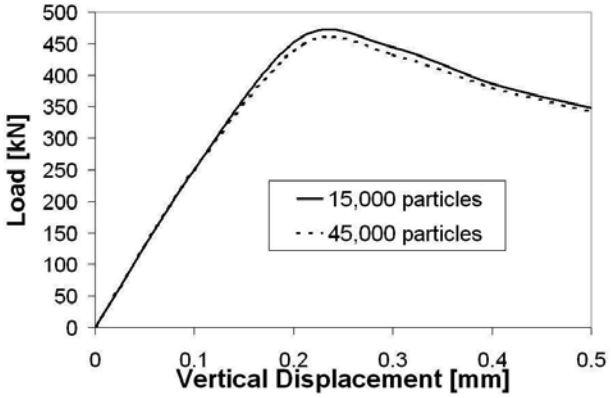
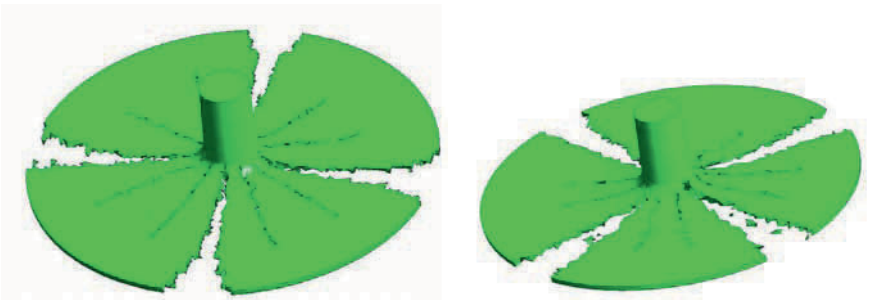


Fig. 5. Load deflection curve of the pull-out test

7.2 Taylor bar impact

To test the method for multiple cracks with crack junction, we consider a Taylor bar impact. There are experimental results available, see Teng et al. (2005). The Taylor bar has a diameter of 6mm and length 30mm. We consider an impact velocities of 600m/s. The failure mechanism is petalling. We have performed similar computations in Rabczuk et al. (2007) but for a much smaller impact velocity where only small strains occur. The material is a 2024-T351 aluminium alloy. We use the Johnson Cook model Johnson and Cook (1983) as in the previous section with Young's modulus $E = 74$GPa, Poisson ratio $\nu = 0.3$, density 0.0027g/mm^3, a reference strain rate of 3.33×10^{-4}, $A = 352$MPa, $B = 440$MPa, $C = 0.0083$, $n = 0.42$, $m = 1$, $c_v = 875$ J/kg C, $T_r = 296$K, $T_m = 775$K and $\beta = 1$. We tested two different discretizations, with approximately 7,000 particles and 22,000 particles, refined adaptively. The final deformation of the Taylor bar is shown in Figure 6 for both discretizations. As can be seen, multiple cracking occurs with overlapping cracks including crack junctions. The failure mode is similar as observed in Teng et al. (2005). There are basically four major cracks that cause the petalling.

Our failure mode looks a little bit ductile, which is likely related to difficulties in the discontinuous bifurcation analysis since the Johnson Cook material does not lose stability easily. The results look almost identical for both discretizations, and agree well with Teng et al. (2005).



a) 22,000 particles b) 7,000 particles

Fig. 6. Final crack pattern of the Taylor bar problem

8 Conclusions

This paper presented a three-dimensional adaptive meshfree method for fracture in statics and dynamics. The initiation, growth, coallescence and branching of an arbitrary number of cracks is handled simply and effectively.

The discontinuities are introduced through extrinsic enrichment of the moving least squares basis, but no near-front enrichment is required. To close the cracks, a Lagrange multiplier field is instead added along the front of the cracks.

Geometrically, the cracks are non-planar surfaces composed of triangular and quadrangular planar sections obtained by cutting the tetrahedral background integration cells by planes whose normals are provided by a material stability analysis.

The results show accurate simulation of large deformation failure problems including fragmentation, where the flexibility of the meshfree method coupled with the efficient crack interaction procedure is most clear. The simulations approach experimental results available in the literature.

References

P.M.A. Areias and T. Belytschko. Analysis of three-dimensional crack initiation and propagation using the extended finite element method. *International Journal for Numerical Methods in Engineering*, 63:760–788, 2005.

Z. P. Bažant and F. C. Caner. Microplane model M5 with kinematic and static constraints for concrete fracture and anelasticity. I: Theory. *ASCE Journal of Engineering Mechanics*, 131(1):31–40, 2005.

T. Belytschko and T. Black. Elastic crack growth in finite elements with minimal remeshing. *International Journal for Numerical Methods in Engineering*, 45(5):601–620, 1999.

T. Belytschko and Y.Y. Lu. Element-free Galerkin methods for static and dynamic fracture. *International Journal of Solids and Structures*, 32:2547–2570, 1995.

T. Belytschko and M. Tabbara. Dynamic fracture using element-free Galerkin methods. *International Journal for Numerical Methods in Engineering*, 39 (6):923–938, 1996.

T. Belytschko, Y.Y. Lu, and L. Gu. Element-free Galerkin methods. *International Journal for Numerical Methods in Engineering*, 37:229–256, 1994.

T. Belytschko, Y.Y. Lu, and L. Gu. Crack propagation by element-free Galerkin methods. *Engineering Fracture Mechanics*, 51(2):295–315, 1995.

T. Belytschko, H. Chen, J. Xu, and G. Zi. Dynamic crack propagation based on loss of hyperbolicity and a new discontinuous enrichment. *International Journal for Numerical Methods in Engineering*, 58(12):1873–1905, 2003.

S. Bordas and B. Moran. Enriched finite elements and level sets for damage tolerance assessment of complex structures. *Engineering Fracture Mechanics*, 73:1176–1201, 2006.

S. Bordas, J.G. Conley, B. Moran, J. Gray, and E. Nichols. A simulation-based design paradigm for complex cast components. *Engineering with Computers*, 2006a. in press, online first.

S. Bordas, V.P. Nguyen, C. Dunant, A. Guidoum, and H. Nguyen-Dang. An extended finite element library. *International Journal for Numerical Methods in Engineering*, 2006b. in press, available online.

G. T. Camacho and M. Ortiz. Computational modeling of impact damage in brittle materials. *International Journal of Solids and Structures*, 33: 2899–2938, 1996.

A. Carpinteri, G. Ferro, and G. Ventura. An augmented lagrangian element-free (ALEF) approach for crack discontinuities. *Computer Methods in Applied Mechanics and Engineering*, 191:941–957, 2001.

J. Dolbow and A. Devan. Enrichment of enhanced assumed strain approximation for representing strong discontinuities: Addressing volumetric incompressibility and the discontinuous path test. *International Journal for Numerical Methods in Engineering*, 59:47–67, 2003.

M. Duflot. A meshless method with enriched weight functions for three-dimensional crack propagation. *International Journal for Numerical Methods in Engineering*, 65(12):1970–2006, 2006.

C. Dunant, S. Bordas, P. Nguyen, A. Guidoum, and H. Nguyen-Dang. Architecture trade-offs of including a mesher in an object-oriented extended finite element code. *European Journal of Computational Mechanics*, 2006. in press.

M. Fleming, Y.A. Chu, B. Moran, and T. Belytschko. Enriched element-free Galerkin methods for crack tip fields. *International Journal for Numerical Methods in Engineering*, 1997.

T.C. Gasser and G.A. Holzapfel. Modeling 3d crack propagation in unreinforced concrete using pufem. *Computer Methods in Applied Mechanics and Engineering*, 2005.

A. Gravouil, N. Moës, and T. Belytschko. Non-planar 3D crack growth by the extended finite element and level sets - part II: Level set update. *International Journal for Numerical Methods in Engineering*, 53:2569–2586, 2002.

G.R. Johnson and W.H. Cook. A constitutive model and data for metals subjected to large strains, high strain rates, and high temperatures. In *Proc. 7th International Symp. on Ballistics*, 1983.

P. Krysl and T. Belytschko. The element free Galerkin method for dynamic propagation of arbitrary 3-D cracks. *International Journal for Numerical Methods in Engineering*, 44(6):767–800, 1999.

Y.Y. Lu, T. Belytschko, and M. Tabbara. Element-free Galerkin method for wave-propagation and dynamic fracture. *Computer Methods in Applied Mechanics and Engineering*, 126(1-2):131–153, 1995.

N. Moës, J. Dolbow, and T. Belytschko. A finite element method for crack growth without remeshing. *International Journal for Numerical Methods in Engineering*, 46(1):133–150, 1999.

N. Moës, A. Gravouil, and T. Belytschko. Non-planar 3-D crack growth by the extended finite element method and level sets, part I: Mechanical model. *International Journal for Numerical Methods in Engineering*, 53(11):2549–2568, 2002.

M. Ortiz and A. Pandolfi. Finite-deformation irreversible cohesive elements for three-dimensional crack-propagation analysis. *International Journal for Numerical Methods in Engineering*, 44:1267–1282, 1999.

T. Rabczuk and T. Belytschko. Adaptivity for structured meshfree particle methods in 2D and 3D. *International Journal for Numerical Methods in Engineering*, 63(11):1559–1582, 2005.

T. Rabczuk and G. Zi. A meshfree method based on the local partition of unity for cohesive cracks. *Computational Mechanics*, 2006.

T. Rabczuk, S. Bordas, and G. Zi. A three dimensional meshfree method for static and dynamic multiple crack nucleation/propagation with crack path continuity. *Computational Mechanics*, 2007. in press, available online.

J. Rethore, A. Gravouil, and A. Combescure. An energy-conserving scheme for dynamic crack growth using the extended finite element method. *International Journal for Numerical Methods in Engineering*, 2005.

X. Teng, T. Wierzbicki, S. Hiermaier, and I. Rohr. Numerical prediction of fracture in the taylor test. *International Journal of Solids and Structures*, 2005.

G. Ventura, J. Xu, and T. Belytschko. A vector level set method and new discontinuity approximations for crack growth by efg. *International Journal for Numerical Methods in Engineering*, 54(6):923–944, 2002.

G. Ventura, E. Budyn, and T. Belytschko. Vector level sets for description of propagating cracks in finite elements. *International Journal for Numerical Methods in Engineering*, 58:1571–1592, 2003.

X.-P. Xu and A. Needleman. Numerical simulations of fast crack growth in brittle solids. *Journal of the Mechanics and Physics of Solids*, 42:1397–1434, 1994.

F. Zhou and J.F. Molinari. Dynamic crack propagation with cohesive elements: a methodolgy to address mesh dependence. *International Journal for Numerical Methods in Engineering*, 59(1):1–24, 2004.

G. Zi and T. Belytschko. New crack-tip elements for xfem and applications to cohesive cracks. *International Journal for Numerical Methods in Engineering*, 57(15):2221–2240, 2003.

G. Zi, J.-H. Song, E. Budyn, S.-H. Lee, and T. Belytschko. A method for growing multiple cracks without remeshing and its application to fatigue crack growth. *Modelling and Simulation in Materials Science and Engineering*, 12(1):901–915, 2004.

G. Zi, H. Chen, J. Xu, and T. Belytschko. The extended finite element method for dynamic fractures. *Shock and Vibration*, 12(1):9–23, 2005.

Accounting for weak discontinuities and moving boundaries in the context of the Natural Element Method and model reduction techniques

F. Chinesta[1], E. Cueto[2], P. Joyot[3], and P. Villon[4]

[1] LMSP UMR 8106 CNRS-ENSAM
151 Boulevard de l'Hôpital, F-75013 Paris, France
francisco.chinesta@paris.esnam.fr
[2] Aragn Institute of Engineering Research, University of Zaragoza. María de Luna, 7. E-50018 Zaragoza, Spain
[3] LIPSI-ESTIA, Technopole Izarbel, F-64210 Bidart, France
[4] Université de Technologies de Compiègne
BP 20529, F-60205 Compiègne cedex, France

Summary. Several thermomechanical models are defined in evolving domains involving fixed and evolving discontinuities. The accurate representation of moving boundaries and interfaces is, despite the significant progresses achieved in the recent years, an active research domain. This work focusses on the application of meshless methods for discretizing this kind of models, and in particular the ones based on the use of natural neighbor interpolations. The questions related to the description of moving boundaries, evolving weak discontinuities and the possibility of an eventual model reduction to alleviate the computational simulation cost, will be some of the topics here analyzed.

Key words: Natural Element Method, Evolving weak discontinuities, Model reduction, Karhunen-Loève, Proper Orthogonal Decomposition, Functional enrichment, Interface tracking and capturing.

1 Introduction

For models involving large transformations the use of meshless discretization techniques seems to be an appealing choice, instead of using the standard finite element method that requires frequent remeshing in order to satisfy the accuracy requirements. Moreover, if one proceeds in the context of meshless methods and all the internal thermomechanical variables are associated to the nodes, neither remeshing nor fields projections are required through the entire

Alain Combescure et al. (eds.), IUTAM Symposium on Discretization Methods for Evolving Discontinuities, 37–55.

simulation. In some cases, addition or deletion of nodes is required in order to capture solution singularities, to improve the solution accuracy or simply to avoid too many nodes in the region where the solution is smooth enough. In these cases, new nodes can be added without complex geometrical checks, in the regions pointed by an adequate error indicator or error estimator, as described for example in [17].

Meshfree methods based on Moving Least Squares (MLS) approximation have been subject of active research during the last decade. These include Smooth Particle Hydrodynamics, Element Free Galerkin, Diffuse Elements, Reproducing Kernel Particle and other Methods (see [1] for a nice review of those techniques). However, one of the issues is the satisfaction of essential boundary conditions. This is due to the nature of the approximation itself. In fact, the MLS nodal domains of influence are the same as those of the corresponding weighting functions, who generally do not fit the boundary. On the other hand, the Natural Neighbor (NN) approximation and associated family of computational methods [12] [3] do not present these drawbacks. The boundary approximation is obtained naturally due to the fact that NN shape functions of internal nodes vanish at the boundary where only the boundary nodes contribute. The list of connected points —the natural neighbors— is also known in advance. However, the NN do not present all the advantages of the MLS. In particular, the shape function support is geometrically complex. Moreover, the NN shape functions have only C^0 continuity at the nodes and only linear consistency is guaranteed. A common difficulty of all these techniques lies in the introduction of discontinuities of the primal variable or of its normal derivative across fixed or moving interfaces as well as the description of moving boundaries as encountered in large transformations of solids (as encountered in forming processes simulation) or in fluid flows.

In what follows, and always in the context of the Natural Element Method, we will consider different possibilities for (i) accounting for the geometrical changes, that is, the accurate representation of complex domain boundary evolutions; and (ii) accounting for fixed or moving weak discontinuities or interfaces (usually encountered in models involving change of phases, consolidation of porous media, ...).

2 The meshfree natural element method

In this section, the utility of both the constrained natural element method (C- NEM) and the -shape based natural element method (α-NEM) to describe moving interfaces and discontinuities in a fixed cloud of nodes is discussed. After a brief review of the Voronoi-based interpolants, we introduce the constrained Voronoi diagram which is used to compute the shape functions in any domain, as well as the α-shapes based approximations functions. To avoid duplication with some of our former publications, different references to our former works will be addressed.

2.1 Natural neighbor interpolation

The construction of the natural neighbor interpolation has been analyzed in depth in some of our former works. The interested reader can refer to [3] for a review on that topic, which we briefly summarize in this section. The NEM interpolant is constructed on the basis of the Voronoi diagram. The Delaunay tessellation is the topological dual of the Voronoi diagram. The Voronoi cells related to neighbor nodes have a common edge.

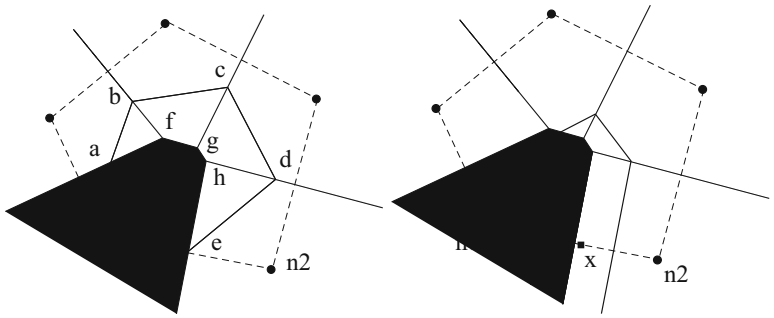


Fig. 1. Construction of the Sibson shape functions.

For the sake of simplicity from now on we focus only on the 2D case, the 3D case being a direct extension. Consider a set of nodes $S = \{n_1, n_2, \ldots, n_N\}$ in $\Re^2$. The Voronoi diagram is the partition of $\Re^2$ into regions T_i (Voronoi cells) defined by:

$$T_i = \{\mathbf{x} \in \Re^2 : d(\mathbf{x}, \mathbf{x}_i) < d(\mathbf{x}, \mathbf{x}_j), \forall j \neq i\}, \quad \forall\ i \tag{1}$$

where $d(\ )$ denotes a distance.

In order to define the natural neighbour coordinates it is necessary to introduce the second-order Voronoi diagram of the cloud defined as

$$T_{ij} = \{\mathbf{x} \in \Re^2 : d(\mathbf{x}, \mathbf{x}_i) < d(\mathbf{x}, \mathbf{x}_j) < d(\mathbf{x}, \mathbf{x}_k) \quad \forall\ j \neq i \neq k\}. \tag{2}$$

Sibson [13] defined the natural neighbor coordinates of a point $\mathbf{x}$ with respect to one of its neighbors n_i as the ratio of the cell T_i that is transferred to T_x when adding $\mathbf{x}$ to the initial cloud of points to the total volume of T_x. In other words, if $\kappa(\mathbf{x})$ and $\kappa_i(\mathbf{x})$ are the Lebesgue measures of T_x and T_{xi} respectively, the natural neighbor coordinates of $\mathbf{x}$ with respect to the node n_i is defined as

$$\phi_i(\mathbf{x}) = \frac{\kappa_i(\mathbf{x})}{\kappa(\mathbf{x})}. \tag{3}$$

Figure 1 illustrates the construction of $\phi_1(\mathbf{x})$, that in this case is given by:

$$\phi_1(\mathbf{x}) = \frac{Area(afghe)}{Area(abcde)} \tag{4}$$

If point $\mathbf{x}$ coincides with node n_i, i.e. $(\mathbf{x} = \mathbf{x}_i)$, $\phi_i(\mathbf{x}_i) = 1$, and all other shape functions are zero, i.e. $\phi_j(\mathbf{x}_i) = \delta_{ij}$ (δ_{ij} being the Kroenecker delta). The properties of positivity, interpolation, and partition of unity are then verified [12]:

$$\begin{cases} 0 \leq \phi_i(\mathbf{x}) \leq 1 \\ \phi_i(\mathbf{x}_j) = \delta_{ij} \\ \sum_{i=1}^{n} \phi_i(\mathbf{x}) = 1 \end{cases} \tag{5}$$

The natural neighbor shape functions also satisfy the local coordinate property [13], namely:

$$\mathbf{x} = \sum_{i=1}^{n} \phi_i(\mathbf{x}) \mathbf{x}_i \tag{6}$$

which combined with Eq. (5), implies that the natural neighbor interpolant spans the space of linear polynomials (linear completeness).

Sibson natural neighbor shape functions are C^1 at any point except at the nodes, where they are only C^0. The C^1 continuity in the domain can be improved by using special classes of natural neighbor shape functions [5], and some ongoing works of Cueto's group allow also to improve the continuity at the nodes by computing B-splines over Voronoi diagrams.

Another important property of this interpolant is its strict linearity over the boundary of convex domains. The proof can be found in Sukumar *et al.* [12]. This result is essential to guarantee strict continuity of the approximation across material interfaces as well as the imposition of essential boundary conditions. The lack of this property is an important issue in most meshfree methods which require special numerical strategies to circumvent this drawback. As just indicated, the property of linearity of the NEM interpolant is only satisfied along convex boundaries [12]. The difficulties related to non-convex geometries can be circumvented using α-shapes [2] or introducing a visibility criterion (C-NEM) [15].

Consider an interpolation scheme for a scalar function $A(\mathbf{x}) : \Omega \subset \Re^2 \rightarrow \Re$, in the form:

$$A^h(\mathbf{x}) = \sum_{i=1}^{n(\mathbf{x})} \phi_i(\mathbf{x}) \; A_i \tag{7}$$

where A_i are the nodal values of the field A at the $n(\mathbf{x})$ neighbor nodes of point $\mathbf{x}$, and $\phi_i(\mathbf{x})$ are the shape functions at that point associated with each neighbor node. It is noted that Eq. (7) defines a local interpolation scheme. Thus, the trial and test functions used in the discretization of a generic variational formulation can be approximated by Eq. (7).

2.2 The constrained natural element method – C-NEM –

In the C-NEM a visibility criterion was introduced in order to restrict influent nodes among natural neighbors [15]. The computation of the shape functions is then done on the basis of the so-called constrained (or extended) Voronoi diagram (CVD) which is the strict dual of the constrained Delaunay triangulation.

The intersection between the constrained Voronoi diagram and the domain closure is composed of cells T_i^C, one for each node n_i, such that any point $\mathbf{x}$ inside T_i^C is closer to n_i than to any other node n_j visible from $\mathbf{x}$. The constrained Voronoi cells are defined formally by:

$$\begin{aligned} T_i^C \;=\; & \{\mathbf{x} \in \Re^n : d(\mathbf{x}, \mathbf{x}_i) < d(\mathbf{x}, \mathbf{x}_j), \\ & \forall j \neq i,\ S_{x \to n_i} \cap \Gamma = \emptyset,\ S_{x \to n_j} \cap \Gamma = \emptyset\} \end{aligned} \tag{8}$$

where Γ is the domain boundary and $S_{a \to b}$ denotes the segment between the points a and b.

A generalization of the constrained Delaunay triangulation to 3D does not exist without adding new nodes. Nevertheless, some techniques for constructing 3D constrained Delaunay tessellations are available and provided in [10] and [11] by addition of Steiner points.

2.3 The α-shape based natural element method – α-NEM –

One important issue when using meshless methods, and particularly when one simulates forming processes from an updated Lagrangian formulation, is free surface tracking. Since, by definition, meshless methods do not need any explicit connectivity between nodes, and consequently the nodes belonging to the domain boundary must be identified with the help of an appropriate technique. In this section we introduce an approach based on the use of the geometrical concept of α-shapes. The concept of shape had traditionally no formal meaning, so it is possible to define a complete family of shapes of a cloud of points by introducing a parameter α that can be considered as a measure of the *level of detail* up to which the domain is going to be represented. α-shapes provide a means so as to eliminate from the triangulation those triangles or tetrahedra whose size is bigger than the before-mentioned level of detail. This criterion is very simple: just eliminate those triangles (tetrahedra) whose circum-radius is bigger than the level of detail, α. For a more in depth description the reader is referred to [2] [3] and the references therein.

In order to clarify the concepts just introduced, we present in the following paragraphs an example of α-shapes computed from a cloud of points corresponding to the simulation of an extrusion process. In this section we will restrict ourselves to geometrical concepts only. The key idea of the method proposed here is to extract the shape of the domain at each time step by invoking the concept of α-shape of the cloud. The α parameter will be obtained by

geometrical considerations. In this case the radius at the outlet of the tooling, for instance, seems to be the smallest level of detail up to which the domain must be represented. The nodal distance h must be chosen accordingly.

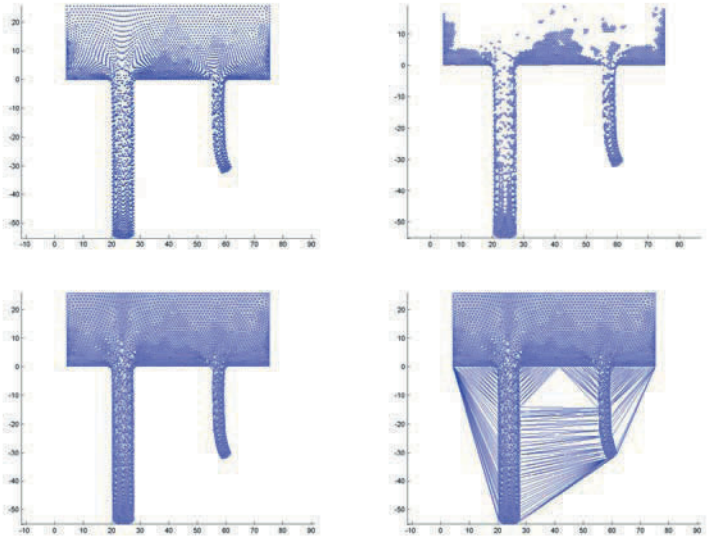


Fig. 2. Some members of the family of α-shapes of the cloud of points used in the extrusion example. (top-left) $\alpha = 0$ (the cloud of points) (top-right) $\alpha \approx 0.5h$ (bottom-left) $\alpha \approx h$ and (bottom-right) $\alpha = \infty$ (the convex hull of the set)

In Fig. 2 some members of the family of α-shapes of the cloud of points in its final configuration are depicted. In Fig. 2(a) the member for $\alpha = 0$, i.e., the cloud of points itself, is shown. Note how, as α is increased, the number and size of the simplexes (in this case, triangles) that belong to the shape is increasing. For $\alpha \approx h$ we obtain an appropriate shape for the cloud. Note, however, that this is not an exact value to be determined at each time step. There exists an interval of acceptable α values for a single shape. Finally, by increasing the α value, the convex hull of the cloud of points is obtained.

This construction allows to reproduce exactly linear polynomials over the boundary of any domain. When dealing with piece-wise homogeneous domains, for instance, it is also necessary to ensure the discontinuity of the derivatives of some field (which is itself continuous across the interface). This can also be done by avoiding natural neighbourhood between nodes placed at both sides of the material interface.

2.4 C-NEM versus α-NEM

The application of the C-NEM requires the definition of the constrained Voronoi diagram, that itself requires an explicit description of the domain boundary, which is usually described from a set of nodes defining a polygonal curve. For this reason this technique seems specially well adapted for treating problems in which the geometry or interfaces evolve moderately [16], or the ones involving fixed or moving cracks [15].

On the contrary the α-NEM seems specially well adapted for treating problems in which the domain geometry evolves significantly, as the ones involving large transformation or complex fluid flows [7]. However its application in the context of fracture mechanics deserves particular treatments, due to the occasional high level of detail induced by the crack separation.

Mixed strategies could be imagined: the α-shapes extracting the domain geometry whereas the interfaces treatment is carried out in the C-NEM framework. This marriage allows to profit the most appealing properties of each one of these strategies.

3 Representation of evolving interfaces

3.1 Tracking versus capturing

To represent evolving weak discontinuities the most standard procedures are the ones based on an interface tracking or capturing it by solving the PDE governing its motion. The first strategy is more simple, but its applicability is restricted to problems in which the interface evolves moderately, because for complex evolutions the tracking algorithm becomes too sophisticated to be efficient. In that follows and for the sake of simplicity we restrict our discussion to the 2D case, however, extension to the 3D case is straightforward.

On the other hand, the interface can be captured using different strategies. In the volume-of-fluid method the interface defines the discontinuity of the characteristic function related to one of the regions that the interface defines. The interface can be defined by the interpolated curve or the elements (when a discontinuous piecewise constant approximation is considered) related to a value of $1/2$ of the characteristic function. This kind of strategies fail when accurate interface descriptions are desired. Another more accurate description defines the interface as the zero value of a level-set function. This level-set function is advected with a certain velocity that on the interface must coincide with the real interface velocity. During its motion, this level-set function degenerates and must be updated frequently to preserve accuracy in the interface description. The most usual correction consists in transforming the advected level set into a signed distance to the interface. Different efficient algorithms exist today for performing both the advection and the updating of that level-set function. The intersection of the curve related to the zero level

set value (extracted by interpolation) with a background mesh can be used as the nodes defining the interface. Now, two possibilities exist from the point of view of the construction of a functional interpolation:

1. To enrich the interpolation to describe the transmission conditions across the interface using Partition of Unity; enforcing reproduction conditions during the construction of the interpolation functions; or by enriching the interpolation in the elements which are intersected by the interface introducing additional degrees of freedom that can be condensed in the original ones as in the cohesive elements framework. If one is applying this last strategy the size of the problem remains constant because no new degrees of freedom are introduced during the simulation.
2. The intersection points between the zero level-set curve and the background Delaunay triangulation can be considered as new discretization nodes. Now, the interpolation can be defined at both interface sides assuring the functional continuity but its discontinuous normal derivative. If the natural element method is been used, the distortion of the Delaunay triangulation in the neighborhood of the interface does not affect to the interpolation accuracy as proved in [16]. However, in this technique the size of the discrete problem is evolving, because the number of intersection points between the interface and the background Delaunay triangulation changes during the motion of the interface.
   If the interface is described by a constant number of points that are advected by the interface velocity (tracking technique), then the size of the discrete problem remains constant, but we can imagine that this possibility can be only envisaged when the interface evolves moderately, because in other cases the addition of new nodes to represent the interface geometry is compulsory.

3.2 Interpolation: enrichment versus explicit interface representation

Interpolation enrichment based on MLS-NEM

There are three main possibilities of enrichment:

1. Partition of unity enrichment. The first one concerns the use of the Partition of Unity paradigm (as usually considered in the framework of the extended finite element – X-FEM – [8]) that generates a linear system whose size evolves as the interface evolves (the nodes of the elements that are intersected by the interface are enriched with new degrees of freedom). As the interpolation generated by the NEM satisfies the partition of unity, all the developments proposed in the context of the finite elements can be extended straightforward to the NEM. This was for example the technique used in [4] for defining mixed interpolations verifying the stability LBB conditions.

2. Element interpolation enrichment. In this case the interpolation in the elements intersected by the interface is enriched by using a function whose derivative becomes discontinuous across the interface and that vanishes at the nodal positions. This new approximation function has associated a new degree of freedom that does not corresponds to a nodal value. By imposing the transmission conditions this new degree of freedom can be written as a function of the nodal values. Thus, this new degrees of freedom are finally condensed on the initial ones, assuring a constant size of the discrete problem which corresponds to the original cloud of nodes.
3. Moving Least Squares enrichment. The consideration of this strategy allows to define an hybrid technique (MLS-NEM) in which different consistencies can be enforced (e.g. the required by material interfaces) without detriment of the appealing NEM properties.

In that follows we focus on the second strategy due to its novelty. Due to the equivalence between the moving least squares and the reproducing kernel particle methods, we are considering by the sake of simplicity the last framework. Let Ω be a 1D domain where the problem is defined (all the results have a direct 2D or 3D counterpart). The points within this domain will be noted by x or s.

The approximation $u^h(x)$ of $u(x)$ is built from the convolution integral

$$u^h(x) = \int_\Omega w(x-s,h)u(s)d\Omega \tag{9}$$

where $w(x-s,h)$ is the kernel function and h a parameter defining the size of the approximation support.

The main idea in the enriched RKPM method [14] is to enforce the reproduction of a general function that we can write in the form of a polynomial plus another function noted by $u^e(x)$:

$$u^h(x) = a_0 + a_1 x + \ldots + a_n x^n + a_{n+1}u^e(x) \tag{10}$$

In the following paragraphs we analyze the required properties of the kernel function $w(x-s,h)$ for reproducing a function expressed by (10).

From Eq. (9), the reproduction of a constant function a_0 is given by

$$\int_\Omega w(x-s,h)a_0 d\Omega = a_0 \tag{11}$$

which implies

$$\int_\Omega w(x-s,h)d\Omega = 1 \tag{12}$$

which constitutes the partition of unity.

Now, the required condition to reproduce a linear function $u^a(x) = a_0 + a_1 x$ is

$$\int_{\Omega} w(x-s,h)(a_0+a_1 s)d\Omega = a_0 + a_1 x \tag{13}$$

By using the partition of unity (12), Eq. (13) can be rewritten as

$$\begin{cases} \int_{\Omega} w(x-s,h)d\Omega = 1 \\ \int_{\Omega} w(x-s,h)sd\Omega = x \end{cases} \tag{14}$$

which implies the linear consistency of the approximation. Repeating this reasoning results

$$\begin{cases} \int_{\Omega} w(x-s,h)d\Omega = 1 \\ \int_{\Omega} w(x-s,h)sd\Omega = x \\ \qquad\qquad \vdots \\ \int_{\Omega} w(x-s,h)s^n d\Omega = x^n \\ \int_{\Omega} w(x-s,h)u^e(s)d\Omega = u^e(x) \end{cases} \tag{15}$$

We will note by $u^r(x)$ the approximation function verifying the conditions (15). Usually a cubic spline is considered as kernel function, and consequently the conditions given by Eq. (15) are not satisfied. Liu et al. [6] propose the introduction of a correction function $C(x, x-s)$ for satisfying the reproduction conditions. In our case we consider the more general form $C(x, s, x-s)$ whose pertinence will be discussed later. Thus $u^r(x)$ will be expressed by

$$u^r(x) = \int_{\Omega} C(x,s,x-s)w(x-s,h)u(s)d\Omega \tag{16}$$

where $C(x, s, x-s)$ is assumed to have the following form

$$C(x,s,x-s) = \mathbf{H}^T(x,s,x-s)\mathbf{b}(x) \tag{17}$$

where $\mathbf{H}^T(x, s, x-s)$ represents the vector containing the functions considered in the approximation basis, and $\mathbf{b}(x)$ is a vector containing unknown functions that will be determined for satisfying the reproduction conditions. Thus, Eq. (15) can be rewritten and the vector $\mathbf{b}(x)$ evaluated after the introduction of a quadrature formula.

Thus, the functional approximation can be expressed as (see [14] for additional details)

$$u^r(x) \cong \sum_{i=1}^{N} \psi_i(x) u_i \tag{18}$$

where ψ_i is the enriched RKP approximation shape function, leading to the so-called enriched reproducing kernel particle approximation (E-RKPA).

To define NN-approximations with discontinuous derivatives we could proceed in the context of the partition of unity (as in the extended finite element technique). However, in this work we propose an enrichment that does not

involve additional degrees of freedom. For this purpose we start introducing the enriched reproducing kernel particle method, that by introducing the NN-interpolation as kernel function leads to NN-interpolation functions with discontinuous derivatives, leading to the so-called enriched natural element interpolation (E-NEM).

To define NN-approximations with discontinuous derivatives we could proceed in the context of the partition of unity (as in the extended finite element technique). However, in this work we propose an enrichment that does not involve additional degrees of freedom. For this purpose we start introducing the enriched reproducing kernel particle method, that by introducing the NN-interpolation as kernel function leads to NN-interpolation functions with discontinuous derivatives.

We consider a level set description $\Theta(\mathbf{x})$ of an interface where the field normal derivatives (with respect to the interface) are discontinuous. Now, we can introduce as enrichment function $u^e(x)$ the following function:

$$u^e(\mathbf{x}) = H_0(\Theta(\mathbf{x}))\Theta(\mathbf{x}) \tag{19}$$

where

$$\Theta(\mathbf{x}) = \begin{cases} \Theta(\mathbf{x}) < 0 \text{ if } \mathbf{x} \in \Omega_1 \\ \Theta(\mathbf{x}) > 0 \text{ if } \mathbf{x} \in \Omega_2 \\ \Theta(\mathbf{x}) = 0 \text{ if } \mathbf{x} \in \Gamma_d \end{cases} \tag{20}$$

and

$$\begin{cases} H_0(\Theta(\mathbf{x})) = 1 \; if \; \Theta(\mathbf{x}) \geq 0 \\ H_0(\Theta(\mathbf{x})) = 0 \; if \; \Theta(\mathbf{x}) < 0 \end{cases} \tag{21}$$

Now, we consider a linear consistency enriched with the function given by Eq. (19) and the kernel function $w(\mathbf{x} - \mathbf{x}_i, h) = \phi_i(\mathbf{x})$ (the natural neighbor shape functions). The resulting approximation shape functions have the linear consistency but allows also to reproduce discontinuous normal derivatives across the interface Γ_d. As the use of the NEM kernel function restricts the number of neighbor nodes to the natural ones, in order to impose higher order consistency new degrees of freedom can be associated to the existing nodes, in a formulation that we called Hermite-NEM, or in other additional nodes, strategy that we called bubble-NEM [18].

To illustrate the capabilities of the proposed technique we consider the exact solution of the Laplace's problem (modelling the temperature distribution in a steady heat transfer problem) defined in a bi-material consisting of two cylinders with different thermal conductivities. The reproduction tests have been carried out using the E-RKPM as well as the E-NEM, where the circular interface was modelled from the distance to that interface that multiplies the Heaviside's function related to that distance. Fig. 3 illustrate a detail of the reconstructed temperature field where we can notice an accurate description of the interface. The discontinuity in the field derivatives is accurately

accounted, as suggested by the representation of the x-derivative depicted in figure 3.

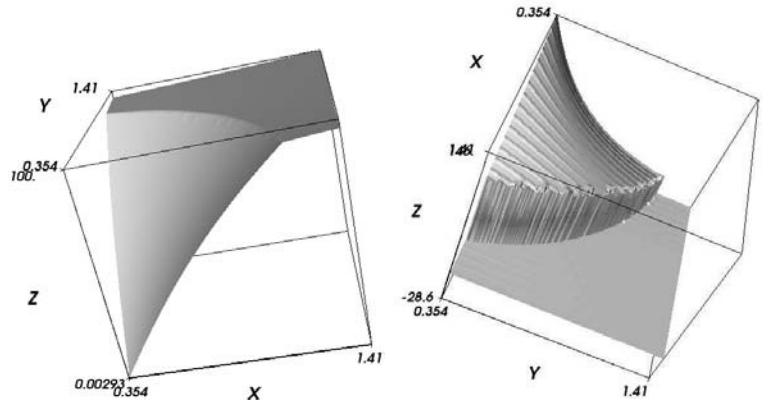


Fig. 3. (left) Enriched Natural Neighbor approximation with discontinuous normal derivatives across a circular interface; (right) x-derivative of the temperature field.

Finally, in order to quantify the accuracy of the results we compare in figure 4 the error (using the two usual norms) using the E-RKPM and the E-NEM techniques. In figure 4(right) we can notice that the E-NEM error is not affected by the slope change across the interface, that increases with the difference of thermal conductivities (for $k1 = 10$ the ratio of conductivities is 10 whereas it is of 100 for $k1 = 100$).

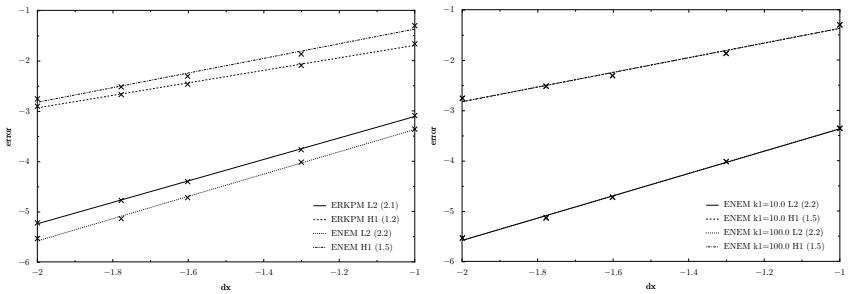


Fig. 4. (left) Approximation errors using the E-RKPM and the E-NEM; (right) E-NEM approximation error for different conductivities ratios.

Explicit interface representation

The ability of the C-NEM for treating problems involving cracks has been illustrated in [15] and for moving interfaces in thermal problems in [16]. In the

present paper, the domain is partitioned in some regions with different material properties. Each subdomain is discretized using a cloud of nodes and the interfaces between the different regions are described by a polygonal curve defined by a set of nodes. Then, a constrained Voronoi diagram is defined at each subdomain with respect to the domain boundary and the interfaces. The attractive feature of the present technique is the possibility to move the interfaces without special care for the shape of the underlying Delaunay triangles because the interpolation accuracy does not depend on the geometrical quality of the Delaunay triangles, in contrast to the FEM. In this manner, the continuity of the approximation is guaranteed by the strict linearity of the interpolation across the interfaces, that are defined by a set of interface nodes.

To illustrate this behavior, we consider the situation depicted in Fig. 5, where the point $\mathbf{x}$ moves from Ω_1 to Ω_2. If $\mathbf{x}$ is in Ω_1, the interpolated field is constructed using the visible neighbor nodes from point $\mathbf{x}$, all of them inside Ω_1 (Γ_I is assumed opaque). If $\mathbf{x}$ is on Γ_I, according to the previous discussion, the interpolated field is strictly linear because it only depends on the two neighbor nodes located on Γ_I. Finally, when $\mathbf{x}$ is in Ω_2, the interpolated field is defined using the visible neighbor nodes from point $\mathbf{x}$ all of them inside Ω_2 (Γ_I being opaque). The continuity of the interpolated field is then guaranteed, but a discontinuity appears in the normal derivative across the interface, because of a sudden change in the neighbor nodes across the interface.

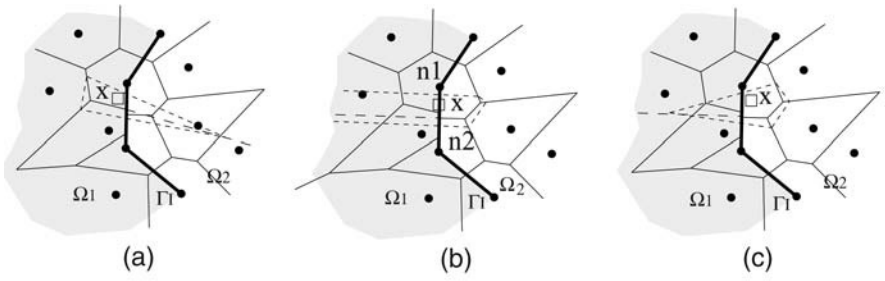


Fig. 5. Reproducing discontinuous derivatives using the constrained Voronoi diagram.

4 Model reduction

4.1 Fundamentals: Karhunen-Loève decomposition and reduced basis enrichment

We assume that the evolution of a certain field $T(\mathbf{x}, t)$ is known. In practical applications, this field is expressed in a discrete form which is known at

the nodes of a spatial mesh and for some times t^m. Thus, we consider that $T(\mathbf{x}_i, t^m) = T^m(\mathbf{x}_i) \equiv T_i^m$ ($t^m = m \times \Delta t$). We can also write $\underline{T}^m$ for the vector containing the nodal degrees of freedom at time t^m. The main idea of the Karhunen-Loève (KL) decomposition is to obtain the most typical or characteristic structure $\phi(\mathbf{x})$ among these $T^m(\mathbf{x})$, $\forall m$. This is equivalent to obtain a function that maximizes α:

$$\alpha = \frac{\sum_{m=1}^{m=M} \left[\sum_{i=1}^{i=N} \phi(\mathbf{x}_i) T^m(\mathbf{x}_i) \right]^2}{\sum_{i=1}^{i=N} (\phi(\mathbf{x}_i))^2} \tag{22}$$

The maximization leads to:

$$\sum_{m=1}^{m=M} \left[\Big(\sum_{i=1}^{i=N} \tilde{\phi}(\mathbf{x}_i) T^m(\mathbf{x}_i) \Big) \Big(\sum_{j=1}^{j=N} \phi(\mathbf{x}_j) T^m(\mathbf{x}_j) \Big) \right] = \alpha \sum_{i=1}^{i=N} \tilde{\phi}(\mathbf{x}_i) \phi(\mathbf{x}_i); \quad \forall \tilde{\phi} \tag{23}$$

which can be rewritten in the form

$$\sum_{i=1}^{i=N} \left\{ \sum_{j=1}^{j=N} \Big[\sum_{m=1}^{m=M} T^m(\mathbf{x}_i) T^m(\mathbf{x}_j) \phi(\mathbf{x}_j) \Big] \tilde{\phi}(\mathbf{x}_i) \right\} = \alpha \sum_{i=1}^{i=N} \tilde{\phi}(\mathbf{x}_i) \phi(\mathbf{x}_i); \ \forall \tilde{\phi} \tag{24}$$

Defining the vector $\underline{\phi}$ such that its i-component is $\phi(\mathbf{x}_i)$, Eq. (24) takes the following matrix form

$$\underline{\tilde{\phi}}^T \, \underline{\underline{c}} \, \underline{\phi} = \alpha \underline{\tilde{\phi}}^T \underline{\phi}; \ \forall \underline{\tilde{\phi}} \Rightarrow \underline{\underline{c}} \, \underline{\phi} = \alpha \underline{\phi} \tag{25}$$

where the two-point correlation matrix is given by

$$c_{ij} = \sum_{m=1}^{m=M} T^m(\mathbf{x}_i) T^m(\mathbf{x}_j) \Leftrightarrow \underline{\underline{c}} = \sum_{m=1}^{m=M} \underline{T}^m (\underline{T}^m)^T \tag{26}$$

which is symmetric and positive definite. If we define the matrix $\underline{\underline{Q}}$ containing the discrete field history:

$$\underline{\underline{Q}} = \begin{pmatrix} T_1^1 & T_1^2 & \cdots & T_1^M \\ T_2^1 & T_2^2 & \cdots & T_2^M \\ \vdots & \vdots & \ddots & \vdots \\ T_N^1 & T_N^2 & \cdots & T_N^M \end{pmatrix} \tag{27}$$

then it is easy to verify that the matrix $\underline{\underline{c}}$ in Eq. (25) results

$$\underline{\underline{c}} = \underline{\underline{Q}} \, \underline{\underline{Q}}^T \tag{28}$$

A posteriori reduced modelling

If some direct simulations have been carried out, we can determine T_i^m, $\forall i \in [1, \cdots, N]$ and $\forall m \in [1, \cdots, M]$, and from these solutions the n eigenvectors related to the n-highest eigenvalues that are expected to contain the most information about the problem solution. For this purpose we solve the eigenvalue problem defined by Eq. (25) retaining all the eigenvalues $\underline{\phi}_k$ belonging to the interval defined by the highest eigenvalue and that value divided by a large enough value (10^8 in our simulations). In practice n is much lower than N. Thus, we can try to use these n eigenfunctions $\underline{\phi}_k$ for approximating the solution of a problem slightly different to the one that has served to define T_i^m. For this purpose we need to define the matrix $\underline{\underline{B}} = [\underline{\phi}_1 \cdots \underline{\phi}_n]$

$$\underline{\underline{B}} = \begin{pmatrix} \phi_1(\mathbf{x}_1) & \phi_2(\mathbf{x}_1) & \cdots \phi_n(\mathbf{x}_1) \\ \phi_1(\mathbf{x}_2) & \phi_2(\mathbf{x}_2) & \cdots \phi_n(\mathbf{x}_2) \\ \vdots & \vdots & \ddots \vdots \\ \phi_1(\mathbf{x}_N) & \phi_2(\mathbf{x}_N) & \cdots \phi_n(\mathbf{x}_N) \end{pmatrix} \tag{29}$$

Now, if we consider the linear system of equations coming from the discretization of a generic problem, in the form:

$$\underline{\underline{G}}\, \underline{T}^m \;=\; \underline{H}^{m-1} \tag{30}$$

where the superscript refers to the time step, then, assuming that the unknown vector contains the nodal degrees of freedom, it can be expressed as:

$$\underline{T}^m = \sum_{i=1}^{i=n} \zeta_i^m \; \underline{\phi}_i \;=\; \underline{\underline{B}}\, \underline{\zeta}^m \tag{31}$$

from which Eq. (30) results

$$\underline{\underline{G}}\, \underline{T}^m \;=\; \underline{H}^{m-1} \Rightarrow \underline{\underline{G}}\, \underline{\underline{B}}\, \underline{\zeta}^m \;=\; \underline{H}^{m-1} \tag{32}$$

and by multiplying both terms by $\underline{\underline{B}}^T$ we obtain

$$\underline{\underline{B}}^T \underline{\underline{G}}\, \underline{\underline{B}}\, \underline{\zeta}^m \;=\; \underline{\underline{B}}^T \underline{H}^{m-1} \tag{33}$$

which proves that the final system of equations is of low order, i.e. the dimension of $\underline{\underline{B}}^T \underline{\underline{G}}\, \underline{\underline{B}}$ is $n \times n$, with $n \ll N$.

Enriching the approximation basis

Obviously, accurate simulations require an error evaluation as well as the possibility of adapting the approximation basis by introducing new functions able to describe the solution features. Ryckelynck proposed in [9] an adaptive procedure, able to construct or enrich the reduced approximation basis. For

this purpose, he proposed to add to the reduced approximation basis some Krylov subspaces generated by the equation residual. Despite the fact that this enrichment tends to increases the number of approximation function, when it is combined with a KL decomposition that continuously reduces this number, the size of the problems is quickly optimized and some times stabilized. This technique is summarized in the present section.

We assume at present that the evolution problem has been solved in the entire time interval using the reduced basis defined by matrix $\underline{\underline{B}}$ (in reference to Eq. (31)) solving Eq. (33):

$$\underline{\zeta}^m \;=\; \left(\underline{\underline{B}}^T \underline{\underline{G}}\; \underline{\underline{B}}\right)^{-1} \underline{\underline{B}}^T \underline{H}^{m-1} \tag{34}$$

We assume that $t_{max} = M \times \Delta t$ and consequently the residual at t_{max}, $\underline{R}^M$, can be computed from

$$\underline{R}^M \;=\; \underline{\underline{G}}\; \underline{\underline{B}}\; \underline{\zeta}^M - \underline{H}^{M-1} \tag{35}$$

If the norm of the residual is small enough $\|\underline{R}^M\| < \epsilon$ (being ϵ a small enough parameter) the computed solution can be assumed as good, but on the contrary, if $\|\underline{R}^M\| \geq \epsilon$, the solution must be improved. For this purpose, we propose to enrich the reduced approximation basis by introducing the just computed residual:

$$\underline{\underline{B}} \leftarrow [\underline{\underline{B}}\; \underline{R}^M] \tag{36}$$

and now, the evolution is recomputed in the entire whole interval using Eq. (34) with the just updated reduced basis $\underline{\underline{B}}$. Both steps, enrichment and the evolution updating, must be repeated until verifying the stop condition $\|\underline{R}^M\| < \epsilon$. After reaching this threshold value, the final reduced approximation basis could be constructed by applying the Karhunen-Loève decomposition to the last time evolution of the solution.

4.2 Accounting for weak discontinuities

When one considers the application of model reduction techniques in problems involving weak discontinuities, two questions arise suddenly: (i) Can the transient solution of problems involving weak discontinuities be expressed as a linear combination os a reduced number of modes?; and (ii) Can the approximation basis be enriched using the residual? The first question determines the reducibility of the problem and the second one the possibility to perform this reduction in a priory approach.

In concluding these two question we consider a one dimensional heat transfer problem defined by:

$$\frac{\partial T_i(x,t)}{\partial t} = k_i \frac{\partial^2 T_i(x,t)}{\partial x^2} + f(x), \qquad x \in \Omega_i \tag{37}$$

where $\Omega_1 =]-1, 0[$ and $\Omega_2 =]0, 1[$, $f(x) = 1$, $k = 1$ and $k = 10$ (all the units in the metric system), being the initial, boundary and transmission conditions given by:

$$\begin{cases} T_1(x = -1, t) = 0 \\ T_1(x, t = 0) = 0 \\ T_2(x = 1, t) = 0 \\ T_2(x, t = 0) = 0 \\ T_1(x = 0, t) = T_2(x = 0, t) \\ -k_1 \left.\frac{\partial T_1}{\partial x}\right|_{x=0,t} = -k_2 \left.\frac{\partial T_2}{\partial x}\right|_{x=0,t} \end{cases} \tag{38}$$

Figure 6 depicts the temperature evolution for $t < 1s$ as well as the functions that results from the Karhunen-Loève decomposition and that are associated to eigenvalues grater than 10^{-8} times the highest one. The reducibility of the model is proved by the existence of only 4 functions from which the entire solution evolution can be expressed. It is easy to prove that the solution of the transient problem expressed in this reduced basis is in perfect agreement with the finite element solution in the entire time interval.

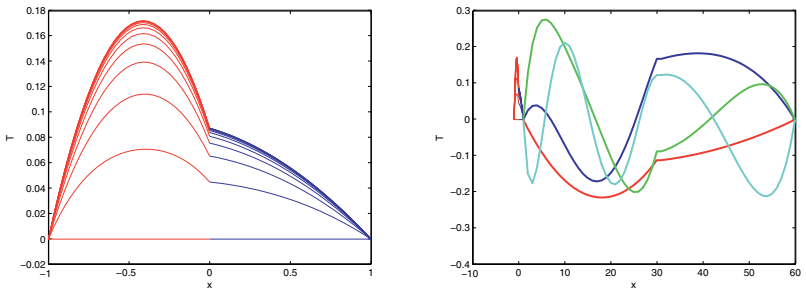


Fig. 6. (left) Time evolution of the temperature profile; (right) functions defining the reduced approximation basis.

To address the second question we consider the solution of the transient problem starting from a reduced basis that only contains the function depicted in figure 7. This reduced basis in enriched by adding the residual computed at time $t = 1$, from which the entire evolution is recomputed. Even if the enrichment increases the size of the discrete problem, the Karhunen-Loève decomposition performed when the convergence is reached allows to reduced the size of the approximation basis. Thus, the size of the discrete problem remains stabilized through the entire simulation.

Figure 7 depicts the function that constitutes the first reduced approximation basis basis as well as the residual associated to the solution at $t = 1$ computed in that reduced basis (broken line). The solution obtained at $t = 1$ when the enrichment algorithm converges is also depicted in this figure, and

we can notice that it corresponds to the one computed using the global finite element basis.

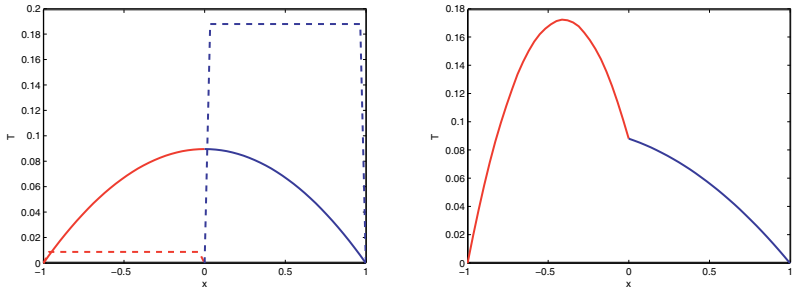


Fig. 7. (left) First approximation function and its associated residual at $t = 1$; (right) Computed solution at $t = 1$ after convergence of the enrichment algorithm.

5 Conclusions

In this paper we have explored some alternatives for treating fixed or evolving weak discontinuities in the context of the meshless natural element method. We have illustrated that standard and new strategies can be applied without detriment of the main appealing properties of this meshless discretization technique.

We have also presented some preliminary results concerning the reduction of such models. In the case of fixed interfaces the reduction procedure works, opening new perspectives in the reduction of models involving evolving discontinuities. This topic that constitutes a work in progress, combines the element enrichment (followed by a static condensation of the new introduced degrees of freedom) and the standard model reduction previously described.

References

1. Belytschko T, Krongauz Y, Organ D, Fleming M, Krysl P(1996) Comp. Meth. in Applied Mechanics and Eng. 139:3–47
2. Cueto E, Doblare M, Gracia L (2000) International Journal for Numerical Methods in Engineering 49:519–546
3. Cueto E, Sukumar N, Calvo B, Cegonino J, Doblare M (2003) Archives of Computational Methods in Engineering: State of the art reviews 10/4:307–384
4. Gonzalez D, Cueto E, Doblare M (2004) International Journal for Numerical Methods in Engineering 61:611-632
5. Hiyoshi H, Sugihara K (2002) Computational Geometry 22:167183.

6. Liu WK, Jun S, Zhang YF (1995) Int. J. Numer. Methods Fluids 21: 1081–1106
7. Martinez MA, Cueto E, Alfaro I, Doblare M, Chinesta F (2004) International Journal for Numerical Methods in Engineering, 60/13: 2105–2129
8. Moes N, Dolbow J, Belytschko T (1999) International Journal for Numerical Methods in Engineering 46: 131-150
9. Ryckelynck D (2005) Journal of Computational Physics 202: 346–366
10. Shewchuck JR (1998) Tetrahedral mesh generation by delaunay refinement. In: Proceedings of the fourteenth annual symposium on computational geometry, Minneapolis, Minnesota. Association for Computing Machinery
11. Shewchuck JR (2000) Sweep algorithms for constructing higher-dimensional constrained Delaunay triangulations. In: Proceedings of the sixteenth annual symposium on computational geometry. Association for Computing Machinery.
12. Sukumar N, Moran B, Belytschko T (1998) International Journal for Numerical Methods in Engineering 43:839–887
13. Sibson R (1980) Math. Proc. Camb. Phil. Soc. 87:151–155
14. Trunzler J, Joyot P, Chinesta F (2005) Discontinuous derivative enrichment in RKPM. In: Lectures Notes on Computational Science and Engineering. Springer 43: 93–108
15. Yvonnet J, Ryckelynck D, Lorong P, Chinesta F (2004) International Journal for Numerical Methods in Engineering 60:1451–1474
16. Yvonnet J, Ryckelynck D, Lorong Ph, Chinesta F (2005) International Journal of Thermal Sciences 44: 559–569
17. Yvonnet J, Ryckelynck D, Coffignal G, Lorong Ph, Chinesta F (2006) Computers & Structures 84/21: 1301–1312
18. Yvonnet J, Villon P, Chinesta F (2006) International Journal for Numerical Methods in Engineering 66/7: 1125–1152

DISCONTINUOUS GALERKIN METHODS

Modeling Evolving Discontinuities with Spacetime Discontinuous Galerkin Methods

Reza Abedi[1], Shuo-Heng Chung[2], Morgan A. Hawker[3], Jayandran Palaniappan[1], and Robert B. Haber[1]

[1] Department of Mechanical Science & Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA; {r-haber, rabedi, palanppn}@uiuc.edu
[2] Department Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA; schung6@uiuc.edu
[3] C&I Engineering, Louisville, KY 40218, USA (formerly, Department of Theoretical & Applied Mechanics, University of Illinois at Urbana-Champaign); mahawker@gmail.com

Summary. We review recent progress in applying spacetime discontinuous Galerkin (SDG) finite element methods to problems whose solutions exhibit various types of moving discontinuities. SDG models and related solution methods offer a number of attractive features, including element-wise satisfaction of the governing balance laws, linear computational complexity in the number of spacetime elements, and a computational structure that readily supports parallel implementations. We describe the use of new unstructured spacetime meshing procedures ind discretizing evolving discontinuities. Specifically, we show how h-adaptive spacetime meshing can be used to capture weak shocks in linear elastodynamics, how the SDG framework provides a convenient setting for implementing cohesive models for dynamic fracture, and how more advanced spacetime meshing procedures can deliver sharp representations of discontinuous solution features by tracking the trajectories of contact discontinuities in compressible gas dynamics.

Key words: spacetime, discontinuous Galerkin, adaptive analysis, cohesive fracture, interface tracking

1 Introduction

1.1 Numerical Representations of Evolving Discontinuities

Discontinuous fields are common in continuum physics, yet their accurate representation remains one of the most challenging problems in computational mechanics, especially when the loci of singular surfaces evolve during the course of a simulation and are unknown *a priori*. Examples of evolving

Alain Combescure et al. (eds.), IUTAM Symposium on Discretization Methods for Evolving Discontinuities, 59–87.

discontinuities include shocks in solid and fluid dynamics, growing cracks and moving phase boundaries arising in heterogeneous flows or the microstructures of solid materials.

Standard conforming finite element models do not admit discontinuous functions. However, even for various nonconforming models that admit discontinuities, singular surfaces are generally restricted to discrete, mesh-related manifolds whose trajectories must be aligned with the discontinuous solution features to obtain an accurate result. This alignment poses special challenges, especially in problems where singular surfaces nucleate spontaneously or alter their connectivity through self-intersection. Beyond simply admitting discontinuous solution features in the discrete model, it is critical to enforce the correct jump conditions arising from the relevant balance laws and to respect the relevant interface kinetics to obtain a physically meaningful model.

Discretizing evolving discontinuities remains a ubiquitous and challenging problem in computational science and engineering and an active area in numerical methods research. We do not attempt a comprehensive review here. Instead, the following subsections identify some of the primary classes of available methods for resolving discontinuous fields, particularly those related to the methods presented in the later sections of this work. Our focus is primarily on finite element methods that address this issue.

Discontinuity/Shock Capturing with *h*-Refinement

We refer to methods that attempt to approximate discontinuous solution fields in continuous solution spaces as *capturing methods*; these are by far the most popular approach. The main advantage of this approach is that no special meshing procedures are needed to make the grid conform to evolving discontinuities, especially when the connectivities of singular surfaces change over time. The mesh might be Lagrangian or Eulerian and structured or unstructured, depending on the problem at hand. The solution space must be sufficiently enriched so that the continuous approximations of jumps are sufficiently sharp. Satisfying this requirement for problems with evolving discontinuities on a static grid can be prohibitively expensive, so some form of adaptive analysis is desirable, even though the mesh is not required to track the discontinuity.

Conforming finite element methods based on Bubnov-Galerkin projections suffer severe numerical artifacts and might even fail, especially for hyperbolic and hyperbolic-parabolic problems, when applied to problems with discontinuous soltuions. Non-local Gibbs oscillations as well as overshoot and undershoot at shocks are common problems. There is an extensive literature on stabilized finite element methods and shock-capturing operators to address these problems. The streamline upwind Petrov-Galerkin (SUPG) method, the Galerkin/least squares (GLS) method and the use of bubble functions exemplify successful finite element techniques for this class of problems (see for example, [8, 9, 23, 24, 26, 42]). The formulation of new methods with improved

shock-capturing properties remains an active research area; recent contributions include spectral vanishing viscosity methods [27, 44] and sub-cell shock capturing [39].

Cohesive Damage Models for Fracture

Cohesive damage models, first introduced by Dugdale [15] and Barenblatt [5], simulate crack initiation and growth by modeling the macroscopic effects of various nonlinear damage processes in the neighborhood of the crack tip. A constitutive *traction–separation law* (TSL) describes the traction acting on a cohesive interface as a nonlinear, bounded function of the separation across the interface [20, 34, 35, 45, 49, 51]. The TSL eliminates the crack-tip stress singularities that arise in classical fracture mechanics and introduces a microscopic length scale that is essential to the fracture model [19].

Numerical implementations of cohesive damage models present three special requirements: the model must admit jumps in the displacement field across cohesive surfaces, there must be a mechanism for enforcing the traction–separation law, and there must be sufficient enrichment to capture the evolution of mechanical fields associated with growing cracks. Special *cohesive elements* embedded in a mesh of conforming finite elements provide a common means to address the first two requirements [10, 13, 49]. The cohesive elements discretize the cohesive interface and the TSL is incorporated into their constitutive model. Alternatively, the cohesive law can also be enforced as a boundary condition on the cohesive interface [34, 45]. More recently, partition of unity, extended finite element or generalized finite element methods have emerged as alternative means to model cracks and cohesive interfaces within finite elements (as opposed to exclusively at boundaries) [6, 14, 31–33, 48]. These methods relax the coupling between the mesh geometry and the crack path.

Sufficient grid refinement is required to capture the detailed response in the vicinity of the active process zone and to enforce accurately the nonlinear traction–separation law. Numerical instabilities can arise if a minimum level of mesh refinement is not realized along the cohesive process zone [20]. The pattern of refinement must evolve to track the moving process zones as a crack propagates. Schrefler et al. periodically remesh the structure to model this phenomenon in quasi-static fracture analysis [41]. In the case of dynamic fracture, mesh refinement must also track the trajectories of sharp wavefronts that are either emitted by moving crack tips or generated by shock loading. Adaptive refinement algorithms can guarantee accuracy and stability, while avoiding the prohibitive expense of uniform mesh refinement. Pandolfi and Ortiz adaptively insert cohesive elements to follow arbitrary crack paths in fragmentation studies [38]. A recent adaptive analysis procedure by Krysl and Belytschko [28] employs an element-free Galerkin method to simplify adaptive refinement and track the crack in an arbitrary direction.

Tracking Methods

We use the term *discontinuity tracking* to describe methods that attempt to model solution discontinuities with true jumps in the discrete solution space by aligning certain mesh features that accommodate jumps with evolving discontinuous features. The most common choice is to associate jumps with element boundaries, but recent applications of extended finite element methods [53] and related partition of unity methods [4] allow greater flexibility by allowing discontinuities to pass through element interiors. However, it is still necessary to align integration cells with intra-element discontinuities and to respect the usual geometric quality constraints on the subdivided element domains.

The mesh must track discontinuities continuously through a time step in dynamics applications, rather than simply update the mesh between time steps. Thus, some form of moving-grid method is required. The Arbitrary Lagrangian–Eulerian (ALE) methods were among the first finite element procedures to implement this strategy [22]; there have been numerous applications since (see for example, [30]). In contrast to capturing methods, tracking methods' support for true jumps in the discrete solution also presents an opportunity to address the associated jump conditions in the formulation.

On the other hand, continuous tracking methods introduce a number of complications. Moving meshes distort element geometries and eventually trigger re-meshing operations to (if constraints on minimum element quality are enforced). Unfortunately, re-meshing itself generates significant errors as the solution is projected from the old mesh to the new. One can limit the mesh distortion by moving free nodes in concert with the motions of nodes constrained to follow a moving singular surface, but this usually adds cost and complexity to the algorithm. In Lagrangian formulations, moving mesh methods generate convective terms that might require special treatment in the finite element formulation. Finally, tracking methods can become intractable when the connectivities of singular surfaces evolve during the course of the simulation. Nonetheless, for problems where they are feasible, their ability to render true discontinuities without strong mesh refinement while enforcing the proper jump conditions often make tracking methods the most accurate and least expensive option for discretizing a discontinuity.

Implicit geometry models, such as level set methods [36], provide a powerful approach for modeling evolving interfaces with complex geometry and varying connectivity (see for example, [17]). These methods combine aspects of capturing and tracking methods; although the level-set contour tracks the evolving interface, a capturing method is typically used (often on a fixed Eulerian grid) to model the discontinuous response.

1.2 Spacetime Discontinuous Galerkin Methods

This contribution reviews recent progress in the use of adaptive spacetime discontinuous Galerkin (SDG) finite element methods for tracking evolving

discontinuities. See [12] for a survey of discontinuous Galerkin finite element methods. In this work, our focus is on spacetime discontinuous Galerkin (SDG) methods that involve direct discretization of spacetime domains and that use basis functions that are fully discontinuous across all spacetime element boundaries. When applied to hyperbolic problems and implemented on suitably constructed spacetime grids, SDG methods support efficient element-by-element or patch-by-patch solution procedures [2, 11, 25, 29, 40, 50]. SDG solution methods share a number of attractive features:

- element-wise satisfaction of the governing balance/conservation laws
- linear computational complexity in the number of spacetime elements (for fixed polynomial order)
- support for high-order approximations on a fixed, compact stencil
- support for spacetime adaptive meshing operations with zero projection error
- a computational structure that lends itself to parallel implementations

The discontinuous nature of the SDG basis functions might appear to address directly the problem of discretizing evolving discontinuities. Indeed, we illustrate the use of this property for discretizing growing cracks and moving contact discontinuities in Subsections 5.2 and 5.3 below. However, these techniques require a careful alignment between the spacetime grid and the trajectories of singular surfaces that might not be easy to achieve. Especially in cases where the connectivity of the discontinuous features changes over time, the flexibility of spacetime adaptive meshing and the accuracy delivered by the element-wise balance properties are of equal or greater importance in capturing moving discontinuities. We demonstrate both approaches below.

1.3 Organization of This Paper

The remaining content of this paper is organized as follows. Section 2 introduces the special notation based on the exterior calculus and differential forms used to formulate our spacetime methods. Section 3 presents summaries of the SDG formulations for linear elastodynamics and for the inviscid Euler equations for gas dynamics that support the numerical examples in this paper. Explanations of our spacetime meshing procedures and the patch-by-patch SDG solution procedure they support appear in Section 4. Section 5 presents three applications of the SDG method to modeling evolving discontinuities. These cover capturing weak shocks in elastodynamics *via* h-adaptive spacetime meshing, implementation of a cohesive damage model for elastodynamic crack growth within the SDG framework, and a study of the use of interface tracking in spacetime for modeling a moving contact discontinuity in a shock tube simulation. Section 6 presents conclusions and prospects for further development.

2 Differential Forms and Notation

This section introduces special notation used in our spacetime formulations that accesses the mathematical machinery of differential forms and the exterior calculus on manifolds. See [7, 18, 43] for a more complete development of these subjects and [2] for more detail on our particular usage. While this notation is non-standard in mechanics it provides a direct coordinate-free notation for the spacetime setting that delivers similar advantages to those offered by direct tensor notation in the purely spatial setting. In particular, we obtain a simple and intuitive means for expressing balance and conservation laws over arbitrary spacetime subdomains that automatically include the relevant jump conditions on the spacetime trajectories of singular surfaces.

Let d be the spatial dimension, and let the reference spacetime analysis domain D be an open $(d+1)$-manifold in $\mathbb{E}^d \times \mathbb{R}$ with a regular boundary. The spacetime coordinates $\left(x^1, \ldots, x^d, t\right)$ are the spatial coordinates followed by the time coordinate and are defined with respect to the ordered basis $(\mathbf{e}_1, \ldots, \mathbf{e}_d, \mathbf{e}_t)$. The dual basis is $\left(\mathbf{e}^1, \ldots, \mathbf{e}^d, \mathbf{e}^t\right)$. We follow the standard summation convention; latin indices range from 1 through d, except the index 't' which denotes time and does not imply summation when repeated.

The top form on D is $\boldsymbol{\Omega} := \mathbf{d}x^1 \wedge \ldots \wedge \mathbf{d}x^d \wedge \mathbf{d}t$, where we make use of the standard ordered basis for 1-forms: $\left(\mathbf{d}x^1, \ldots, \mathbf{d}x^d, \mathbf{d}t\right)$. The standard ordered basis for d-forms is $(\mathbf{d}\hat{x}_1, \ldots, \mathbf{d}\hat{x}_d, \star\mathbf{d}t)$, in which $\mathbf{d}\hat{x}_j := \star\mathbf{d}x^j$ where $\star$ is the Hodge star operator. Thus, $\mathbf{d}x^j \wedge \mathbf{d}\hat{x}_k = \delta^j{}_k \boldsymbol{\Omega}$; $\mathbf{d}t \wedge \mathbf{d}\hat{x}_k = \mathbf{0}$; $\mathbf{d}t \wedge \star\mathbf{d}t = \boldsymbol{\Omega}$; and $\mathbf{d}x^j \wedge \star\mathbf{d}t = \mathbf{0}$. We identify $\mathbf{d}\hat{x}_j$ and $\star\mathbf{d}t$ with, respectively, the differential spacetime volume element that is the geometric dual of $\mathbf{e}^j$ and the purely spatial differential volume element. We use $\boldsymbol{i} := \boldsymbol{i}_{\mathbf{e}_t}$ to denote the temporal insertion operator.

We use forms with scalar and tensor coefficients to develop our theory. Bold Italic symbols denote forms (with either scalar or tensor coefficients), while bold non-Italic symbols denote vector and tensor fields on D. We introduce a special 1-form with vector coefficients, $\boldsymbol{d}\mathbf{x} := \mathbf{e}_i \boldsymbol{d}x^i$, and a special d-form with vector coefficients, $\star\boldsymbol{d}\mathbf{x} := \mathbf{e}^i \boldsymbol{d}\hat{x}_i$. We define the exterior product of two forms with tensor coefficients as

$$\mathbf{a}\boldsymbol{\psi} \wedge \mathbf{b}\boldsymbol{\omega} := \mathbf{a}(\mathbf{b})(\boldsymbol{\psi} \wedge \boldsymbol{\omega}) \tag{1}$$

in which $\mathbf{a}$ and $\mathbf{b}$ are tensor fields on D of orders m and n respectively $(m \geq n)$, and $\boldsymbol{\psi}$ and $\boldsymbol{\omega}$ are p and q-forms on D such that $p + q \leq d + 1$.

3 Formulations

This section presents brief reviews of the SDG formulations for linearized elastodynamics and the Euler equations for inviscid gas dynamics. Component expansions are provided to illuminate the relationship between the differential forms expression. More detailed formulations can be found in the references cited below.

3.1 Elastodynamics

This subsection presents a summary of a single-field SDG formulation of linearized elastodynamics [2]. The displacement field $\mathbf{u} = u_i \mathbf{e}^i$, the primary solution variable on D, is related to the strain and velocity 1-forms according to

$$\boldsymbol{E} = \mathbf{E} \wedge \boldsymbol{d}\mathbf{x} = E_{ij} \mathbf{e}^i \otimes \mathbf{e}^j \wedge \mathbf{e}_k \boldsymbol{d}x^k = E_{ij} \mathbf{e}^i \boldsymbol{d}x^j \tag{2}$$

$$\boldsymbol{v} = \dot{\mathbf{u}} \boldsymbol{d}t = \dot{u}_i \mathbf{e}^i \boldsymbol{d}t \tag{3}$$

where $\mathbf{E} = \operatorname{sym} \nabla \mathbf{u}$. These two kinematic quantities sum to form a single spacetime 1-form $\boldsymbol{\varepsilon}$ called the *strain-velocity*:

$$\boldsymbol{\varepsilon} := \boldsymbol{E} + \boldsymbol{v} \tag{4}$$

The fundamental force-like quantities are the *stress* and *linear momentum density* d-forms given by

$$\boldsymbol{\sigma} := \bar{\boldsymbol{\sigma}} \wedge \star \boldsymbol{d}\mathbf{x} = \sigma^{ij} \mathbf{e}_i \otimes \mathbf{e}_j \wedge \mathbf{e}^k \boldsymbol{d}\hat{x}_k = \sigma^{ij} \mathbf{e}_i \boldsymbol{d}\hat{x}_j \tag{5}$$

$$\boldsymbol{p} := \mathbf{p} \star \boldsymbol{d}t = \rho \dot{u}^i \mathbf{e}_i \star \boldsymbol{d}t \tag{6}$$

where $\bar{\boldsymbol{\sigma}}$ is the symmetric Cauchy stress tensor, $\mathbf{p}$ is the linear momentum density vector and ρ is the mass density. The standard relationship of linear elasticity between $\boldsymbol{\varepsilon}$ and $\bar{\boldsymbol{\sigma}}$ is strongly enforced. The restriction of $\boldsymbol{\sigma}$ to a spacetime d-manifold Γ with arbitrary orientation delivers the surface traction $\mathbf{t}_\Gamma$ acting on Γ: $\boldsymbol{\sigma}|_\Gamma = \mathbf{t}_\Gamma$.

The *spacetime momentum flux* $\boldsymbol{M}$ is defined as

$$\boldsymbol{M} := \boldsymbol{\sigma} - \boldsymbol{p} \tag{7}$$

such that the restriction of $\boldsymbol{M}$ to Γ delivers the flux of linear momentum across Γ. The exterior derivative of $\boldsymbol{M}$ is the $(d+1)$-form whose vector coefficient is the residual of the homogeneous Equation of Motion:

$$\boldsymbol{d}\boldsymbol{M} = (\boldsymbol{\nabla} \cdot \bar{\boldsymbol{\sigma}} - \dot{\mathbf{p}}) \, \boldsymbol{\Omega} = \left(\bar{\sigma}^{ij}{}_{,j} - \rho \ddot{u}^i \right) \mathbf{e}_i \boldsymbol{\Omega} \tag{8}$$

The *body force* is given by the $(d+1)$-form, $\boldsymbol{b} = \mathbf{b}\boldsymbol{\Omega} = b^i \mathbf{e}_i \boldsymbol{\Omega}$, where $\mathbf{b}$ is the body force vector per unit mass. Thus, the vector coefficient of the $(d+1)$-form $\boldsymbol{d}\boldsymbol{M} + \rho \boldsymbol{b}$ is the residual of the Equation of Motion.

Consider the integral form of Balance of Linear Momentum,

$$\int_{\partial Q} \boldsymbol{M} + \int_Q \rho \boldsymbol{b} = \mathbf{0} \quad \forall Q \subset D \Rightarrow \int_Q (\boldsymbol{d}\boldsymbol{M} + \rho \boldsymbol{b}) = \mathbf{0} \quad \forall Q \subset D \tag{9}$$

in which subdomains Q are assumed to have regular boundaries, and we use the Stokes Theorem to eliminate the boundary flux integral. Let Γ^J be the union of jump manifolds on D where $\boldsymbol{M}$ and $\boldsymbol{\varepsilon}$ are possibly discontinuous.

Apply the Localization Theorem to (9) to obtain the Equation of Motion, including the jump part of $\boldsymbol{dM}$, denoted $[\boldsymbol{M}]$, on $\Gamma^J \cup \partial D$:

$$(\boldsymbol{dM} + \rho\boldsymbol{b})|_{D\setminus\Gamma^{\mathrm{J}}} = \mathbf{0} \tag{10a}$$

$$[\boldsymbol{M}]|_{\Gamma^{\mathrm{J}}\cup\partial D} = \mathbf{0} \tag{10b}$$

We replace (10b) with a stronger condition,

$$(\boldsymbol{M}^* - \boldsymbol{M})|_{\Gamma^{\mathrm{J}}\cup\partial D} = \mathbf{0} \tag{11}$$

in which $\boldsymbol{M}^*$ is either a Godunov or prescribed value of the momentum flux.

In the SDG setting, we introduce a mesh partition $\mathcal{P}$ of D into finite elements, and introduce a discrete bases so that $\boldsymbol{M}$ and $\boldsymbol{\varepsilon}$ are continuous within each element but are possibly discontinuous across element boundaries. Then the strong form of balance of linear momentum augmented by a kinematic compatibility condition can be expressed as follows. Let V^Q be the solution space for the displacement field on element $Q \in \mathcal{P}$. We then have the following strong form and weighted residual statement:

Problem 1 (Strong Form). For each $Q \in \mathcal{P}$, find $\mathbf{u} \in V^Q$ such that

$$\boldsymbol{dM} + \rho\boldsymbol{b} = \mathbf{0} \quad \text{in } Q \tag{12a}$$

$$(\boldsymbol{M}^* - \boldsymbol{M})|_{\partial Q} = \mathbf{0} \tag{12b}$$

$$(\boldsymbol{\varepsilon}^* - \boldsymbol{\varepsilon})|_{\partial Q} = \mathbf{0} \tag{12c}$$

$$(\mathbf{u}_0^* - \mathbf{u}_0)\star\boldsymbol{dt}|_{\partial Q^{\mathrm{ti}}} = \mathbf{0} \tag{12d}$$

in which $\mathbf{u}_0$ is a projection of the displacement solution $\mathbf{u}$ into a subspace with vanishing total energy, ∂Q^{ti} is the time-inflow boundary of Q and the values $\boldsymbol{M}^*$, $\boldsymbol{\varepsilon}^*$ and $\mathbf{u}_0^*$ are either Godunov or prescribed values, as described below. The jump conditions (12c,d) enforce kinematic boundary conditions and compatibility across element boundaries.

Problem 2 (Weighted Residual Statement). For each $Q \in \mathcal{P}$, find $\mathbf{u} \in V^Q$ such that $\forall\hat{\mathbf{u}} \in V^Q$,

$$\int_Q \dot{\hat{\mathbf{u}}} \wedge (\boldsymbol{dM} + \rho\boldsymbol{b}) + \int_{\partial Q} \left\{\dot{\hat{\mathbf{u}}} \wedge (\boldsymbol{M}^* - \boldsymbol{M}) + (\boldsymbol{\varepsilon}^* - \boldsymbol{\varepsilon}) \wedge \boldsymbol{i}\hat{\boldsymbol{\sigma}}\right\} + \int_{\partial Q^{\mathrm{ti}}} k\hat{\mathbf{u}} \wedge (\mathbf{u}_0^* - \mathbf{u}_0)\star\boldsymbol{dt} = 0 \tag{13}$$

in which $\hat{\mathbf{u}}$ is the weighting function, and k is a constant introduced for dimensional consistency.

An application of the Stokes Theorem leads to the weak problem statement:

Problem 3 (Weak Form). For each $Q \in \mathcal{P}$, find $\mathbf{u} \in V^Q$ such that $\forall\hat{\mathbf{u}} \in V^Q$,

$$\int_Q (-\boldsymbol{d}\dot{\hat{\mathbf{u}}} \wedge \boldsymbol{M} + \dot{\hat{\mathbf{u}}} \wedge \rho \boldsymbol{b}) + \int_{\partial Q} \left\{ \dot{\hat{\mathbf{u}}} \wedge \boldsymbol{M}^* + (\boldsymbol{\varepsilon}^* - \boldsymbol{\varepsilon}) \wedge \boldsymbol{i}\hat{\boldsymbol{\sigma}} \right\} + \int_{\partial Q^{\mathrm{ti}}} k\hat{\mathbf{u}} \wedge (\mathbf{u}_0^* - \mathbf{u}_0) \star \boldsymbol{d}t = 0 \tag{14}$$

Replacing V^Q in Problem 3 with a discrete subspace, $V_h^Q \subset V_Q$, generates the SDG finite element formulation. It is easy to show that the discrete form of Problem 3 balances both linear and angular momentum over every spacetime element [2].

We define the target boundary values of $\boldsymbol{M}$ and $\boldsymbol{\varepsilon}$ for each $Q \subset D$:

$$\boldsymbol{M}^* = \begin{cases} \boldsymbol{M} & \text{on } \partial Q^{\mathrm{co}} \cup (\partial Q \cap \partial D^{\boldsymbol{\varepsilon}}) \\ \boldsymbol{M}^+ & \text{on } \partial Q^{\mathrm{ci}} \setminus \partial D^{\mathrm{ci}} \\ \boldsymbol{M}^{\mathrm{G}}(\boldsymbol{M}, \boldsymbol{M}^+, \boldsymbol{a}_{\partial Q}) & \text{on } \partial Q^{\mathrm{nc}} \setminus \partial D^{\mathrm{nc}} \\ \underline{\boldsymbol{M}} & \text{on } \partial Q \cap \left(\partial D^{\mathrm{ci}} \cup \partial D^{\boldsymbol{M}}\right) \end{cases} \tag{15a}$$

$$\boldsymbol{\varepsilon}^* = \begin{cases} \boldsymbol{\varepsilon} & \text{on } \partial Q^{\mathrm{co}} \cup \left(\partial Q \cap \partial D^{\mathbf{M}}\right) \\ \boldsymbol{\varepsilon}^+ & \text{on } \partial Q^{\mathrm{ci}} \setminus \partial D^{\mathrm{ci}} \\ \boldsymbol{\varepsilon}^{\mathrm{G}}(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^+, \mathbf{a}_{\partial Q}) & \text{on } \partial Q^{\mathrm{nc}} \setminus \partial D^{\mathrm{nc}} \\ \underline{\boldsymbol{\varepsilon}} & \text{on } \partial Q \cap \left(\partial D^{\mathrm{ci}} \cup \partial D^{\boldsymbol{\varepsilon}}\right) \end{cases} \tag{15b}$$

in which $\boldsymbol{M}$ and $\boldsymbol{\varepsilon}$ are traces of interior fields on Q, $\boldsymbol{M}^+$ and $\boldsymbol{\varepsilon}^+$ are traces of exterior fields on Q, ∂Q^{co} is the causal (space-like) outflow boundary of Q, ∂Q^{ci} (∂D^{ci}) is the causal inflow boundary of Q (D), and ∂Q^{nc} (∂D^{nc}) is the non-causal (time-like) boundary of Q (D). The non-causal domain boundary ∂D^{nc} is partitioned into disjoint parts, $\partial D^{\boldsymbol{M}}$ and $\partial D^{\boldsymbol{\varepsilon}}$, where $\boldsymbol{M}$ and $\boldsymbol{\varepsilon}$ are determined by the prescribed functions $\underline{\boldsymbol{M}}$ and $\underline{\boldsymbol{\varepsilon}}$. The Godunov values, $\boldsymbol{M}^{\mathrm{G}}$ and $\boldsymbol{\varepsilon}^{\mathrm{G}}$, are obtained from the solutions to local Riemann problems [2].

3.2 Euler Equations

We define the conservation fields: mass density ρ, linear momentum density $\mathbf{p} = \rho\mathbf{v}$, and total energy denstiy $\mathcal{E} = e + \frac{1}{2}|\mathbf{v}|^2$ where e is the specific internal energy. Let $p = \frac{\rho e}{\gamma - 1}$ be the pressure with adiabatic index γ and $\bar{\boldsymbol{\sigma}} = p\delta^{ij}\mathbf{e}_i \otimes \mathbf{e}_j$ be the stress. The *spacetime conservation fluxes* and their exterior derivatives are

$$\boldsymbol{F}_\rho = \rho \star \boldsymbol{d}t + \rho\mathbf{v} \wedge \star \boldsymbol{d}\mathbf{x} \quad \boldsymbol{d}\boldsymbol{F}_\rho = (\dot{\rho} + \boldsymbol{\nabla} \cdot \rho\mathbf{v})\, \boldsymbol{\Omega} \tag{16a}$$

$$\boldsymbol{F}_{\mathbf{p}} = \mathbf{p} \star \boldsymbol{d}t + (\mathbf{p} \otimes \mathbf{v} - \bar{\boldsymbol{\sigma}}) \wedge \star \boldsymbol{d}\mathbf{x} \quad \boldsymbol{d}\boldsymbol{F}_{\mathbf{p}} = [\dot{\mathbf{p}} + \boldsymbol{\nabla} \cdot (\mathbf{p} \otimes \mathbf{v} - \bar{\boldsymbol{\sigma}})]\, \boldsymbol{\Omega} \tag{16b}$$

$$\boldsymbol{F}_{\mathcal{E}} = \mathcal{E} \star \boldsymbol{d}t + [\mathcal{E}\mathbf{v} + \bar{\boldsymbol{\sigma}}(\mathbf{v})] \wedge \star \boldsymbol{d}\mathbf{x} \quad \boldsymbol{d}\boldsymbol{F}_{\mathcal{E}} = \left\{ \dot{\mathcal{E}} + \boldsymbol{\nabla} \cdot [\mathcal{E}\mathbf{v} + \bar{\boldsymbol{\sigma}}(\mathbf{v})] \right\} \boldsymbol{\Omega} \tag{16c}$$

Note that the exterior derivatives render local residual forms of conservation of mass, linear momentum and energy. We rewrite the fluxes in (16) in vector format as

$$\boldsymbol{F} := \left\{ \begin{array}{l} \boldsymbol{F}_\rho(\mathbf{q}) \\ \boldsymbol{F}_\mathbf{p}(\mathbf{q}) \\ \boldsymbol{F}_\mathcal{E}(\mathbf{q}) \end{array} \right\} \qquad \mathbf{q} := \left\{ \begin{array}{l} \rho \\ \mathbf{p} \\ \mathcal{E} \end{array} \right\}. \tag{17}$$

Following arguments similar to those in the previous subsection, we write the integral and local forms of the system conservation statement $\forall Q \subset D$:

$$\int_{\partial Q} \boldsymbol{F} = \mathbf{0} \Leftrightarrow \left\{ \begin{array}{l} \boldsymbol{dF}|_Q = \mathbf{0} \\ [\boldsymbol{F}]|_{\partial Q} = (\boldsymbol{F}^* - \boldsymbol{F})|_{\partial Q} = \mathbf{0}. \end{array} \right. \tag{18}$$

We use Godunov values for the fluxes $\boldsymbol{F}^*$; this ensures satisfaction of the entropy inequality over each subdomain Q.[4] The jump conditions in (18) are the *Rankine-Hugoniot* conditions that govern the motions of shocks.

We introduce a discrete subspace V_h^Q and apply a simple Bubnov-Galerkin projection to develop the SDG weighted residual statement and the weak form used for numerical implementation.

Problem 4 (SDG Weighted Residual Statement). For each $Q \in \mathcal{P}(D)$, find $\mathbf{q} \in V_h^Q$ such that

$$\int_Q \hat{\mathbf{q}} \wedge \boldsymbol{dF} + \int_{\partial Q} \hat{\mathbf{q}} \wedge (\boldsymbol{F}^* - \boldsymbol{F}) = \mathbf{0} \; \forall \hat{\mathbf{q}} \in V_h^Q. \tag{19}$$

Problem 5 (SDG Weak Statement). For each $Q \in \mathcal{P}(D)$, find $\mathbf{q} \in V_h^Q$ such that

$$-\int_Q \boldsymbol{d}\hat{\mathbf{q}} \wedge \boldsymbol{F} + \int_{\partial Q} \hat{\mathbf{q}} \wedge \boldsymbol{F}^* = \mathbf{0} \; \forall \hat{\mathbf{q}} \in V_h^Q \tag{20}$$

A piece-wise constant choice for $\hat{\mathbf{q}}$ in (20) proves element-wise conservation:

$$\int_{\partial Q} \boldsymbol{F}^* = \mathbf{0} \; \forall Q \in \mathcal{P}(D) \tag{21}$$

In general, stabilization is required to control local overshoot and to ensure a robust numerical method. However, in problems with convex flux functions and where a limited amount of overshoot is acceptable, the basic method presented here can be applied.

4 Causal Spacetime Meshing and Patch-wise Solution Procedure

The construction of causal spacetime meshes and a patch-by-patch solution procedure with linear computational complexity in the number of patches, as described in [1, 2, 37], are key aspects of our implementation of the SDG

[4] This has been proven only for scalar case in [37].

method. We briefly summarize these concepts here; the reader is referred to [1,2,37] for a more detailed treatment.

We use an advancing-front mesh generation procedure called *Tent Pitcher* [16,46] to construct spacetime meshes in 2D × time that satisfy a patch-wise *causality condition*. A patch is a collection of contiguous tetrahedral spacetime elements whose boundary is, therefore, a collection of spacetime triangles. A triangle is *causal* if all the characteristics of the governing hyperbolic equations have the same orientation relative to the triangle at all points on the triangle. A patch is causal if all its boundary faces are either causal or part of the domain boundary. This implies that each boundary face of a causal patch can be classified as either an inflow, outflow or boundary face depending on the face's location and whether the characteristic flow across the face is inward or outward relative to the patch. Since no information can enter a patch through its outflow faces, the solution in each patch depends only on characteristic data flowing through its inflow and boundary faces. This implies a dependency graph that defines a partial ordering of patches wherein the solution on each patch depends only on boundary data and solutions on earlier patches in the partial ordering. Thus, patches can be solved locally, without approximation, using only boundary data and outflow data from previously-solved patches. Inter-element boundaries within a patch can be noncausal, so the elements within a patch must, in general, be solved simultaneously.

Tent Pitcher begins with a constant-time triangulation of the spatial domain at the initial time of the analysis interval. It then visits each vertex in the triangulation in turn to construct a new patch. The basic tent pitching procedure involves advancing a vertex in time to construct a new "tent pole" and then constructing a set of tetrahedra surrounding the tent pole to form a new patch. The height (duration) of the tent pole is limited to ensure that the outflow faces of the new patch are all causal. The tent-pole height might be further limited by a *progress constraint* to ensures that it is always possible to continue to pitch new patches. The product of the tent-pitching algorithm is an unstructured, patch-wise causal spacetime mesh. The local nature of the causality condition only enforces a local restriction on the duration of individual patches, similar to a CFL condition; there is no global restriction on the time step size, as in conventional explicit time-marching schemes. Thus, patches containing elements with larger spatial diameters and slower wave speeds can have longer durations than patches with smaller spatial diameters and higher wave speeds.

Our solution procedure interleaves mesh generation with the finite element solution; we compute the finite element solution on each patch as soon as it is generated by Tent Pitcher. Since the number of tetrahedral elements within a patch is limited by the highest vertex degree in the space triangulation, and because the highest polynomial order in each element is bounded, we can show that this patch-by-patch solution procedure has linear complexity in the number of spacetime patches. There is no need to assemble and store a global system of equations; the finite element routines are written to solve a single

patch at a time in a serial implementation. In a parallel computation, on the other hand, multiple tents can be pitched and solved simultaneously on separate processors, subject only to the partial ordering constraint for patches. The local character of the algorithm reduces the need for interprocessor communication.

5 Applications

This section presents three example applications that demonstrate different ways SDG finite element models can be used to model solutions with evolving discontinuities.

5.1 Discontinuity Capturing Using Adaptive Spacetime Meshing

We next present a brief review of how h-adaptive spacetime meshing can be used to capture weak shocks in elastodynamics, as reported in [3].[5] We seek local refinement along the trajectories of shocks, so that refinement in one location does not impose a global restriction on the durations of elements elsewhere in the spacetime analysis domain. Although the SDG formulation for elastodynamics presented in Subsection 3.1 balances linear and angular momentum over every spacetime element to within machine precision, the method is dissipative, and energy is generally not balanced. Experience shows that an error measure based on element-wise dissipation is an effective indicator for adaptive refinement, both for controlling the overall solution accuracy and for concentrating refinement along the trajectories of shocks. To achieve an efficient solution, we attempt to distribute a limited amount of numerical dissipation evenly over the spacetime elements.

Error Indicator for Element-wise Numerical Dissipation

The numerical energy dissipation for spacetime element Q is given by:

$$\varphi^Q = \frac{1}{2}\int_{\partial Q} (\dot{\mathbf{u}}^* \wedge \boldsymbol{M}^* + \boldsymbol{\varepsilon}^* \wedge \mathbf{i}\boldsymbol{\sigma}^*) + \int_Q \dot{\mathbf{u}} \wedge \rho \boldsymbol{b}. \tag{22}$$

Let φ^* be the user-specified target dissipation per element. The dissipation on element Q is considered acceptable when $\underline{\varphi} \leq \varphi^{\mathrm{Q}} \leq \overline{\varphi}$, where $\underline{\varphi} = (1-\eta)\varphi^*$ and $\overline{\varphi} = (1+\eta)\varphi^*$ in which η is a user-specified parameter subject to $0 < \eta < 1$. Element Q is marked for refinement when $\varphi^Q > \overline{\varphi}$, and element Q is coarsenable when $\varphi^Q < \underline{\varphi}$. Otherwise, element Q is marked acceptable. The parameter η must be chosen sufficiently large to minimize undesirable cycling between coarsening and refinement. We use $\eta = 0.2$ in our current implementation.

[5] More precisely, the features in this example are sharp wavefronts that approximate weak shocks. In the interest of conciseness, we nonetheless refer to these solution features as shocks.

Adaptive Tent Pitcher Algorithm

We use an adaptive extension of the Tent Pitcher algorithm to implement adaptive refinement and coarsening within our patch-by-patch, advancing-front solution algorithm. Rather than adapting patches of spacetime elements directly, Tent Pitcher implements adaptive refinement by managing the triangulation of the current front. Each time a patch is solved, the solver checks whether any elements in the patch have been marked for refinement. If the result is positive, then the dissipation error is deemed unacceptable, and the solver *rejects* the patch when it is returned to Tent Pitcher. Tent Pitcher, in turn, discards the rejected patch and, using a newest-vertex-bisection algorithm [1], refines the triangles on the current mesh front that correspond to the elements marked for refinement. This effectively refines the subsequent spacetime mesh in both space and time when tent pitching is resumed, because the causality constraint dictates shorter tent-pole heights (local time steps) at vertices associated with refined triangles. Note that Tent Pitcher discards only the solution on the rejected patch. The solutions on all previously-solved patches are unaffected due to the patch-wise causal structure of the spacetime grid, so the amount of redundant calculation due to refinement is small.

Tent Pitcher accepts the solution on the current patch if all elements in the patch are either acceptable or coarsenable. In this case, Tent Pitcher stores the patch solution, advances the mesh front, and copies the status (acceptable or coarsenable) from the patch elements to the corresponding facets of the new mesh front. Requests for coarsening need not be acted on immediately, since they do not involve unacceptable error. Tent Pitcher's coarsening operation involves deleting a degree-4 vertex so as to merge two pairs of adjacent, coarsenable triangles into two two triangles in the active mesh front. A vertex can be deleted when all of the triangles surrounding it are marked coarsenable. In order to maintain the integrity of the spacetime grid, each pair of triangles must be coplanar before the coarsening operation can be executed. Typically, this requirement is not satisfied immediately, so Tent Pitcher postpones coarsening until it has pitched new tent poles with heights adjusted to meet the coplanarity constraint.

Numerical Example: Crack-Tip Wave Scattering

We next review results reported in greater detail in [2] that demonstrate the ability of the SDG formulation with adaptive spacetime meshing to capture shock-like features in elastodynamics. Figure 1 shows a center-cracked plate that we model using plane-stress assumptions, Young's modulus $E = 10$, Poisson ratio $\nu = 0.3$, and density $\rho = 2$. A spatially uniform tensile traction of magnitude T acts along the top and bottom edges of the plate. The magnitude T ramps rapidly from zero at time $t = 0.0$ to a maximum value of 10 at time $t = 0.002$ to approximate a weak stress-velocity shock; T holds constant at the maximum value until the simulation interval terminates at time $t =$

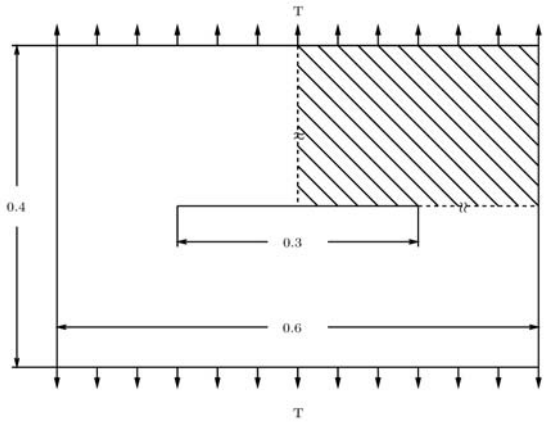


Fig. 1. Spatial domain and boundary conditions for the crack scattering problem

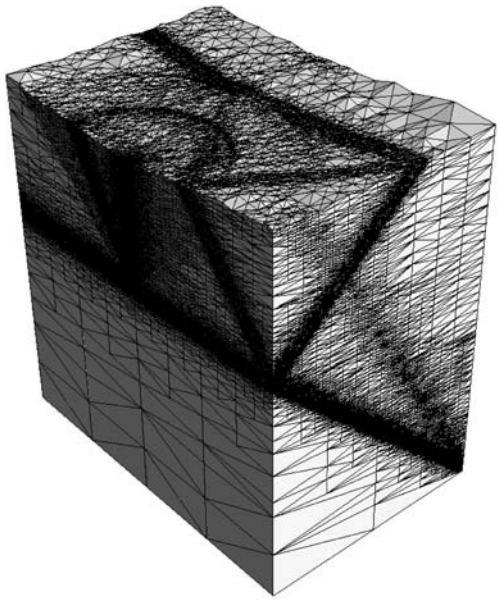


Fig. 2. Adaptive spacetime mesh for crack-tip wave scattering problem with 11 million tetrahedra (intermediate stage, front is at roughly $t = 0.100$)

0.300. We enforce symmetry boundary conditions to model only the shaded region shown in the figure, and we use complete cubic polynomials to model the displacement field within each spacetime tetrahedron. A uniform 2×4 rectangular grid defines the initial space mesh, with each rectangle subdivided into two triangles.

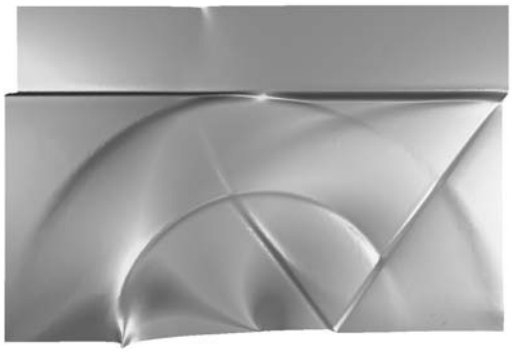


Fig. 3. Visualization of elastic wave scattering by a stationary crack-tip at $t = 0.105$)

Fig. 2 shows the state of the spacetime mesh constructed by the adaptive Tent Pitcher algorithm at an intermediate stage of the simulation. The spatial directions are aligned with the horizontal axes, and time increases upward in the vertical direction. The fine details of the elastodynamic solution are clearly evident in the pattern of mesh refinement. The dark diagonal bands along the right face of the spacetime volume are traces of plane-wave shocks generated by the sudden traction loading, while the lighter diagonal band ascending at a steeper angle traces the trajectory of a Rayleigh wave moving along the free edge of the plate. The spacetime trajectory of the crack tip runs along the vertical center-line of the left face of the spacetime volume. The apex of the cone-shaped region of mesh refinement indicates the initial scattering event. The separation of the scattered wave into dilatational and shear components can be seen in the pattern of refinement. The outer perimeter of the cone (tangent to the dark band on the top surface due to the plane wave) indicates the progress of the faster-moving dilatational wave, while the dark circular band within the cone traces the trajectory of the slower shear wave. In this example, the ratio of the largest to smallest element diameter is 1024. The ability of the adaptive SDG method to limit refinement to the trajectories of moving shocks and the absence of global time-step size constraints yield significant computational savings.

Fig. 3 shows a visualization of the adaptive solution at roughly the same time as the top of the spacetime volume in Fig. 2. The images were generated by a pixel-exact visualization system [52]. The strain energy density field is mapped to color (rendered as gray scale here); the velocity magnitude is mapped to the height field, which is then shaded by a lighting model to reveal its form. The visualization reveals fine features of the solution such as a Rayleigh wave moving along the crack surface and the scattered shear and pressure waves emanating from the crack-tip. Small-scale features, such as the shock fronts, are well resolved. The solution is free of spurious oscillations,

although no extra stabilization was added to the SDG formulation for this linear problem.

5.2 Cohesive Fracture Model

Cohesive damage models are a popular means for modeling discontinuous deformation fields in computational solid mechanics. First introduced by Dugdale [15] and Barenblatt [5], these relatively simple models describe fracture and delamination processes, including an initiation criterion and a means to calculate the rate of crack extension. They require supplementary criteria to determine the direction of crack growth. The works of Needleman and Xu [34,49] led to a strong resurgence of interest in cohesive damage models; the recent literature is extensive (see for example, [6, 10, 13, 14, 28, 31–33, 35, 48]). We do not attempt a comprehensive literature review here. Rather, we describe an implementation of the Xu and Needleman traction separation law for elastodynamic fracture [49] within the adaptive SDG framework described in the previous subsection. In this context, we note the related work of Huang and Costanzo [21].

By using cohesive damage models, one attempts to simulate crack initiation and extension by modeling the macroscopic effects of various nonlinear damage processes in the neighborhood of the crack tip. Specifically, a constitutive *traction–separation law* (TSL) describes the tractions acting on a cohesive interface as a nonlinear, bounded function of the interface separation. The TSL eliminates the crack-tip stress singularities that arise in classical fracture mechanics and introduces a microscopic length scale that is essential to the fracture model [19]. We restrict our attention to two spatial dimensions ($d = 2$) and to the history-independent, exponential relationship developed by Xu and Needleman [49], although the computational framework we describe could be adapted to most other TSLs.

Incorporation of Cohesive Damage Model in SDG Framework

In general, numerical implementations of cohesive damage models present three special requirements: the numerical model must admit jumps in the displacement field across cohesive interfaces; there must be a mechanism for enforcing the traction–separation law; and there must be some form of adaptive enrichment to resolve the active cohesive process zone and to capture the evolution of mechanical fields associated with growing cracks, including the sharp wavefronts generated by shock loads and by sudden crack initiation or arrest. Recalling the discontinuous format of the SDG basis, we address the first requirement by aligning element boundaries with cohesive interfaces. We weakly enforce the traction–separation law by introducing a special definition of the target momentum flux $\boldsymbol{M}^*$ on cohesive interfaces, as described below. The adaptive refinement strategy described in the preceding subsection effectively addresses the third requirement. However, we describe below the use

of an additional error indicator to ensure that the SDG solution renders the TSL accurately.

Let Γ^{C} denote the union of all the cohesive-interface trajectories in $\overline{D}$, and modify the partition of the non-causal domain boundary according to $\partial D^{\mathrm{nc}} = \overline{\partial D^{\mathbf{M}} \cup \partial D^{\boldsymbol{\varepsilon}} \cup \partial D^{\mathrm{C}}}$, where $\partial D^{\mathrm{C}} = \partial D \cap \Gamma^{\mathrm{C}}$. The only modification to the elastodynamic formulation required to implement the cohesive model is to replace (15) with

$$\boldsymbol{M}^* = \begin{cases} \boldsymbol{M} & \text{on } \partial Q^{\mathrm{co}} \cup (\partial Q \cap \partial D^{\boldsymbol{\varepsilon}}) \\ \boldsymbol{M}^+ & \text{on } \partial Q^{\mathrm{ci}} \setminus \partial D^{\mathrm{ci}} \\ \boldsymbol{M}^{\mathrm{G}}(\boldsymbol{M}, \boldsymbol{M}^+, \boldsymbol{a}_{\partial Q}) & \text{on } \partial Q^{\mathrm{nc}} \setminus \left(\partial D^{\mathrm{nc}} \cup \Gamma^{\mathbf{C}}\right) \\ \underline{\boldsymbol{M}} & \text{on } \partial Q \cap \left(\partial D^{\mathrm{ci}} \cup \partial D^{\boldsymbol{M}}\right) \\ \boldsymbol{M}^{\mathrm{C}} & \text{on } \partial Q \cap \Gamma^{\mathrm{C}} \end{cases} \tag{23a}$$

$$\boldsymbol{\varepsilon}^* = \begin{cases} \boldsymbol{\varepsilon} & \text{on } \partial Q^{\mathrm{co}} \cup \left[\partial Q \cap \left(\partial D^{\mathbf{M}} \cup \Gamma^{\mathbf{C}}\right)\right] \\ \boldsymbol{\varepsilon}^+ & \text{on } \partial Q^{\mathrm{ci}} \setminus \partial D^{\mathrm{ci}} \\ \boldsymbol{\varepsilon}^{\mathrm{G}}(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^+, \mathbf{a}_{\partial Q}) & \text{on } \partial Q^{\mathrm{nc}} \setminus \left(\partial D^{\mathrm{nc}} \cup \Gamma^{\mathbf{C}}\right) \\ \underline{\boldsymbol{\varepsilon}} & \text{on } \partial Q \cap \left(\partial D^{\mathrm{ci}} \cup \partial D^{\boldsymbol{\varepsilon}}\right) \end{cases} \tag{23b}$$

in which $\boldsymbol{M}^{\mathrm{C}} := \tilde{\mathbf{s}} \boldsymbol{dx}^{\mathrm{C}}$, where $\boldsymbol{dx}^{\mathrm{C}}$ is the spacetime d-volume element on Γ^{C} and $\tilde{\mathbf{s}}$ is the cohesive traction vector generated by the traction–separation relation.

We use two adaptive error indicators to ensure accurate solutions for both the bulk and cohesive responses. The first error indicator is the element-wise numerical dissipation, as described above. The second indicator uses the L-2 norm of the traction error along the cohesive interface. We use this indicator to ensure that the finite element tractions match the cohesive traction-separation law to within one percent. The combination of these adaptive criteria captures sharp wave-fronts in the bulk and strong gradients in the crack-tip fields.

Adaptive Enforcement of the Traction–Separation Relation

Cohesive models can produce numerical instability when too few elements are included in the active cohesive process zone [20]. Beyond stability problems, under-resolved finite element grids can generate significant errors in the work of separation and the history of crack-tip motion. Adaptive control of the numerical dissipation does not directly address these problems, so we introduce a second adaptive error indicator to limit the mismatch between the tractions generated by the finite element stress solution and the target cohesive tractions.

The *cohesive traction error on* Γ^{C} is given by $\tilde{\mathbf{s}} - \mathbf{s}$, where $\mathbf{s}$ is the vector traction coefficient of $\boldsymbol{\sigma}|_{\Gamma^{\mathrm{C}}}$. We define the *cohesive traction error indicator on element* Q as

$$\tau^Q := \frac{\| |\tilde{\mathbf{s}} - \mathbf{s}| \|_{L_2(\partial Q \cap \Gamma^{\mathrm{C}})}}{\|\bar{s}\|_{L_2(\partial Q \cap \Gamma^{\mathrm{C}})}} \tag{24}$$

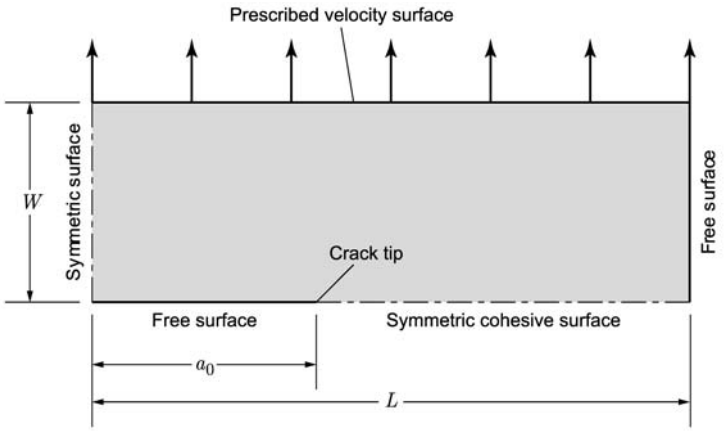


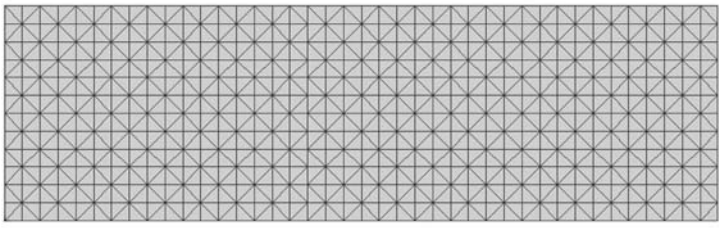
Fig. 4. Model for middle-crack tension specimen.

in which $\bar{s}$ is the cohesive strength parameter in the Xu and Needleman model. Let τ^* be the target value of the cohesive traction error indicator. The cohesive traction error on element Q is acceptable when $\underline{\tau} \leq \tau^{\mathrm{Q}} \leq \overline{\tau}$, where $\underline{\tau} = (1-\eta)\tau^*$ and $\overline{\tau} = (1+\eta)\tau^*$ in which $0 < \eta < 1$. Refinement is required when $\tau^{\mathrm{Q}} > \overline{\tau}$, and element Q might be coarsenable (depending on whether it is also coarsenable with respect to the dissipation error) when $\tau^{\mathrm{Q}} < \underline{\tau}$.

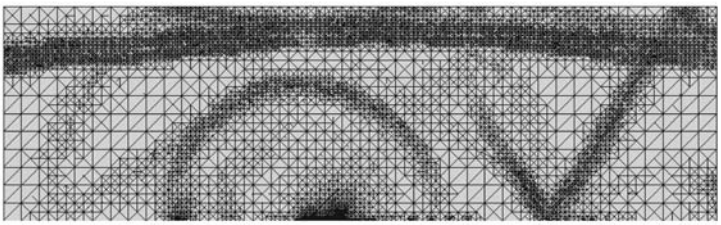
Numerical Example: Cohesive Fracture Under Shock Loading Conditions

We consider cohesive crack growth in a middle-crack tension specimen loaded by a uniform, prescribed velocity along two opposite edges. We use symmetry conditions to model a single quadrant with dimensions, $L = 10$ mm by $W = 3$ mm, and an initial crack length, $a_0 = 4.25$ mm, as shown in Figure 4. The bulk material properties and the parameters for the traction–separation law are identical to those used in [49]. A uniform, prescribed velocity, applied along the top edge of the specimen, ramps from zero to a sustained velocity of 15 m/s over an interval of 0.1 μs. This approximates an elastodynamic shock that reaches the crack tip at $t = 1.44$ μs. We use tetrahedral spacetime elements with complete cubic polynomial bases.

The adaptive SDG model's ability to resolve shocks accurately and to maintain the fidelity of the numerical approximation of the TSL is key to the integrity of this study. Figure 5a shows the spatial discretization that initiates the adaptive solution process. Both adaptive error indicators are active; the target per-element energy dissipation is $\varphi^* = 5 \times 10^{-17}$ J, and the target normalized cohesive traction error is $\tau^* = 1\%$. Figure 5b shows the pattern of adaptive mesh refinement at roughly $t = 3.5$ μs, well after the initial, stationary tip scatters the main shock wave and the crack begins



(a) Initial space mesh



(b) Refinement pattern on mesh front as crack starts to propagate

Fig. 5. Progression of mesh refinement during solution.

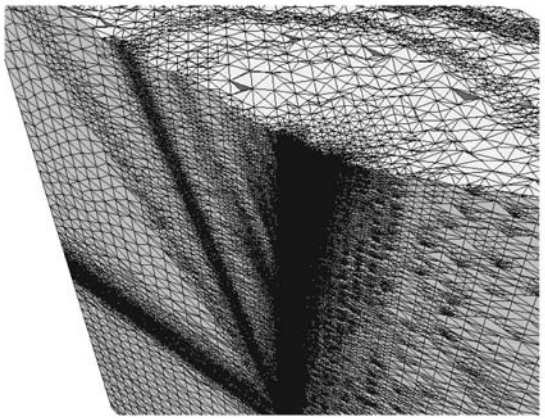


Fig. 6. Detail of spacetime mesh in vicinity of crack-growth-initiation event

to propagate. The pattern of mesh refinement reflects the pattern of shock fronts in the bulk material. The apparent distortion of the wavefronts is due to the fact that the spacetime mesh does not advance uniformly in time; it is not due to solution error. There is a region of refinement along the cohesive interface near the initial tip position where the adaptive solution tracks the moving cohesive process zone. There are strong gradients in the stress field surrounding the crack tip, but the TSL ensures that the cohesive tractions remain bounded.

Figure 6 shows a detail of the portion of the spacetime mesh (the vertical axis represents time) that captures the transtion from a stationary to a running crack. The diagonal bands visible on the front face of the mesh are associated with, from left to right, the dilatational and shear waves scattered by the stationary tip followed by a series of waves emitted by the moving cohesive process zone as it accelerates to the right.

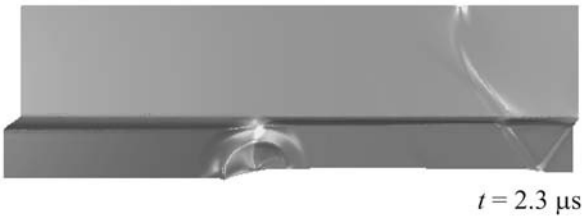
Figure 7 presents three still images from an animation of the SDG solution generated by the per-pixel-accurate spacetime rendering procedure described in [52]; the sequence depicts the transition from a stationary crack to a propagating cohesive failure. The color field (rendered in gray scale here) represents the log of the strain-energy density; the height field depicts the modulus of the material velocity.

Figure 7a shows wave scattering shortly after the shock reaches the crack plane; a sharp gradient in the energy density is evident at the crack tip. The cohesive process zone is stationary, and the response is similar to that of a stationary, mathematically sharp crack. However, as expected, the cohesive model eliminates the singular stress field at the tip. Figure 7b shows the transition to crack propagation. A spike begins to develop in the velocity field near the initial crack-tip position. Post-simulation analysis shows that the velocity field eventually develops the same $r^{-\frac{1}{2}}$ singular structure predicted by linear elastodynamic solutions for running cracks [19], where r is the radial distance to the core of the singularity. The singularity grows in strength and moves with the accelerating crack tip, as shown in Figure 7c. The emergence of the velocity singularity was unexpected, and to our knowledge, has not been predicted by other numerical models for this problem. Our solutions show that the velocity singularity persists through two decades of the radius r. However, in view of the discrete nature of our solutions, it is not clear whether the singular structure persists to arbitrarily small radii. A future publication will present a more detailed investigation of the velocity singularity for running cracks.

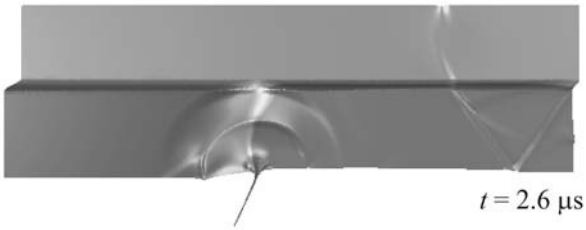
There are also qualitative differences in the crack-growth kinetics predicted by the SDG solution and the Xu and Needleman model. Initiation at 1.6 μs is nearly instantaneous upon the arrival of the primary shock front in [49], and the crack-tip undergoes rapid acceleration to approach the Rayleigh wave speed. The SDG solution, on the other hand, shows a lag of roughly 1 μs between the arrival of the primary shock front and crack-growth initiation, and the subsequent crack-tip acceleration is more gradual. The slower response in the SDG model might be explained by the finite time required to transition from the non-singular velocity field surrounding the stationary crack to the singular form generated by the running crack.

5.3 Front Tracking and Combined Tracking/Capturing Methods

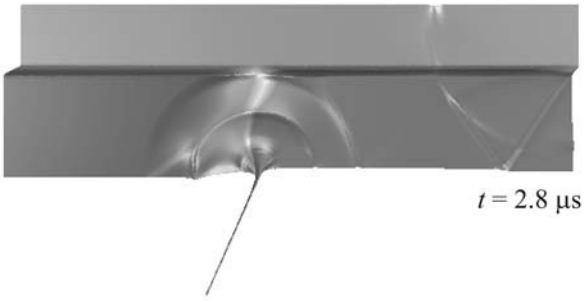
This section presents an example in which we combine two distinct methods for resolving discontinuities. After introducing an expanded set of adaptive



(a) Wave scattering by stationary crack; growth is delayed



(b) Singular velocity field develops as crack growth initiates



(c) Singularity strengthens as crack tip accelerates

Fig. 7. Visualization of crack propagation over time.

meshing operations, we describe a shock-capturing technique for nonlinear conservation laws as well as an interface tracking method that delivers sharp resolution of a moving contact discontinuity by aligning the spacetime mesh with the trajectory of the singular surface.

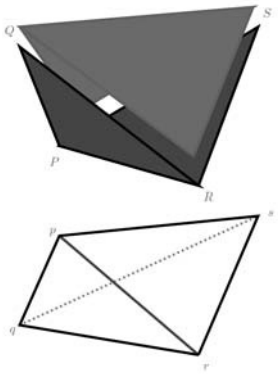


Fig. 8. Edge-flip operation implemented in spacetime (top) vs. instantaneously (bottom)

New Adaptive Meshing Operations

The adaptive meshing procedures used in the preceding examples, although effective, relied on a limited number of meshing operations. We performed mesh refinement instantaneously by subdividing triangles in the space-like mesh front according to a newest vertex scheme. Mesh coarsening required a preliminary step wherein a patch of triangles to be coarsened must first be pitched to a coplanar spacetime configuration prior to each vertex deletion. No provision was made for the spacetime analogues of edge-flips and vertex smoothing — operations that maintain and improve mesh quality in conventional adaptive algorithms. An expanded, more robust set of spacetime adaptive meshing procedures in the current implementation of Tent Pitcher improves solution accuracy and provides new capabilities for tracking discontinuous and non-smooth solution features.

In the previous implementation, the "tent poles" generated by advancing a vertex in time were constrained to strictly vertical orientations (parallel to the time coordinate axis). The current implementation accommodates inclined tent poles, such that the spatial position of a vertex can vary continuously as it advances in time, while ensuring that all patches satisfy the causality and progress contraints required by our $\mathcal{O}(N)$ solution scheme. We use this new capability to carry out smoothing operations on the spatial projection of the mesh front to improve and maintain element quality. That is, we compute a smoothed spatial position for each unconstrained vertex in the space mesh as an area-weighted average of the centroids of the surrounding triangles. An inclined tent pole then moves the vertex from its previous position to the smoothed location. Smoothing is applied every time an unconstrained vertex is pitched, so good mesh quality is maintained throughout the spacetime meshing process. Inclined tent poles are also important in the interface-tracking method described below.

All adaptive operations other than mesh refinement,[6] including edge flips, smoothing and vertex deletions, are now implemented over spacetime patches in a manner that eliminates projection errors associated with remeshing. For example, we use the spacetime edge-flip operation, depicted in Fig. 8, to maintain the Delaunay property in the spatial projection of the mesh front. In the spacetime version, the dark gray inflow triangles, PQR and PRS, of the spacetime tetrahedron $PQRS$ coincide with outflow faces of previously solved elements, so no projection error is incurred. Thus, the edge flip from PR to QS, can be executed with zero projection error. If this operation is carried out instantaneously, as in conventional remeshing procedures, there is unavoidable projection error incurred by switching from piecewise polynomials defined over triangles pqr and prs to polynomials defined on triangles qsp and qrs. Overall, the new set of spacetime meshing operations increases solution accuracy while reducing the number of elements in an adapted spacetime mesh.

Tracking Moving Interfaces

Discontinuous Galerkin methods in which the mesh explicitly tracks evolving discontinuities present an attractive alternative, where feasible, to discontinuity-capturing methods. If element boundaries are aligned with the trajectories of singular surfaces, then the SDG basis functions can model discontinuous solutions directly, including explicit treatment of the relevant jump conditions, without the expensive mesh refinement required in capturing methods. Further, the need for extra stabilization in the vicinity of a discontinuity can be reduced or eliminated. The expanded set of spacetime meshing operations provides a framework for a new approach to interface tracking. We use inclined tent poles to generate patches with internal element boundaries that track the singular surface. For example, we can use a zero-mass-flux condition to determine the proper tent-pole inclination to track a material interface in a heterogeneous flow, or we can use a constutive kinetics equation to track the motion of a phase boundary in a solid. Adaptive smoothing, edge-flip, refinement and vertex-deletion operations maintain the integrity and resolution of the mesh surrounding the moving interface. All these operations are local and are carried out continuously as the simulation proceeds. We thus avoid the mesh distortion and the projection errors associated with global remeshing operations in many other interface tracking methods.

Numerical Example: Combined Interface Tracking and Shock Capturing

Figs. 9 and 10 illustrate a shock-tube simulation based on the inviscid Euler equations (cf. Subsection 3.2) in which we use a combination of techniques

[6] Mesh refinement is still carried out instantaneously on the mesh front. However, this is not problematic because refinement does not induce any projection error and does not require a preliminary synchronization step.

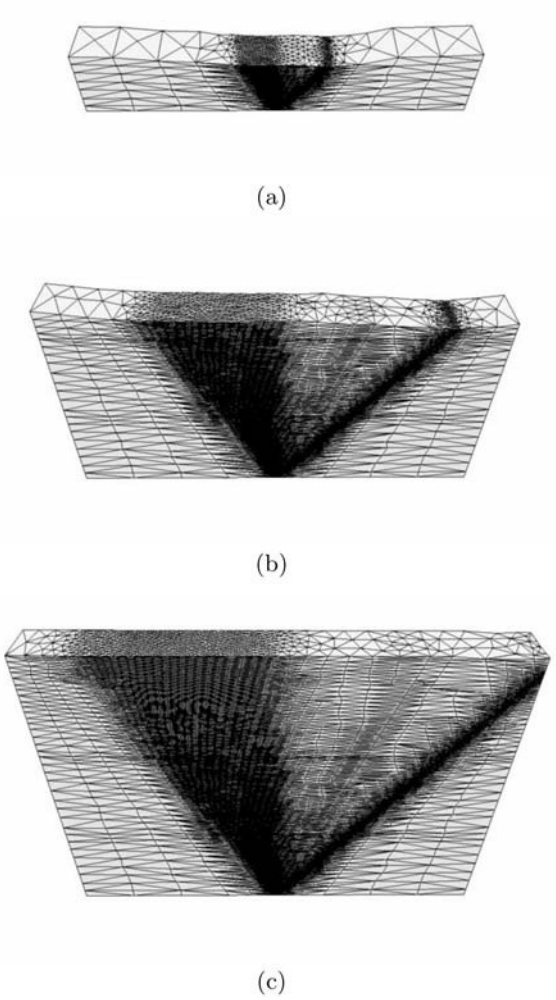


Fig. 9. Shock tube example. Adaptive spacetime mesh tracks trajectory of contact discontinuity moving to right at half the shock speed

to model the discontinuous solution features. Although the basic SDG formulation features some intrinsic upwinding, the nonlinear structure of this problem requires some additional stabilization to control overshoot and undershoot. We use uniform artificial viscosity and piecewise linear polynomials in this preliminary study. Better resolution of the shocks would be possible using a shock-capturing scheme and higher-order polynomials. However, since testing the SDG approach to interface tracking is the main objective of this study, we defer these improvements to future work. We use adaptive spacetime refinement to capture a shock moving from the domain center to the

Fig. 10. SDG solution for shock tube example; height and color fields (rendered in gray scale here) indicate density and pressure. Interface tracking delivers sharp resolution of contact discontinuity.

right and an expansion wave opening to the left. Regions of higher density in the spacetime meshes shown in Fig. 9 identify each feature.

Three residual-based error indicators (one for each conservation equation) in each spacetime element Q drive this adaptive simulation. The error indicators share a common form, given by

$$e_i = \frac{\|\mathbf{dF}_i\|_{L_1(Q)} + \|\mathbf{F}_i - \mathbf{F}_i^*\|_{L_1(\partial Q)}}{\|\rho_i^0\|_{L_1(\Omega^0)}}; \quad i = 1, 2, 3 \tag{25}$$

where Ω^0 is the spatial analysis domain at the initial time and

$$\mathbf{F}_1 := \mathbf{F}_\rho \quad \rho_1^0 = \rho^0 \tag{26a}$$

$$\mathbf{F}_2 := \mathbf{F}_\mathbf{p} \quad \rho_2^0 = \sqrt{\rho^0 \mathcal{E}^0} \tag{26b}$$

$$\mathbf{F}_3 := \mathbf{F}_\mathcal{E} \quad \rho_3^0 = \mathcal{E}^0 \tag{26c}$$

in which ρ^0 and $\mathcal{E}^0$ are the mass density and total energy density fields on Ω^0. Adaptive refinement is triggered when any of the three error indicators exceeds its maximum allowable value in any element in a patch; adaptive coarsening is triggered when all three error indicators fall below a minimum threshold in all the elements in a patch. As seen in Fig. 9, this approach effectively limits spacetime mesh refinement to the trajectories of the shock and the expansion wave. However, as with any capturing method and especially due to the crude form of the stabilization used here, the shock appears smeared out in Fig. 10. Nonetheless, the conservation properties of the SDG model ensures that the shock moves at the correct velocity.

We use a different approach, *interface tracking*, to resolve the contact discontinuity in the density field. This feature appears as a coherent surface in the spacetime mesh midway between the center of the domain and the right-traveling shock; it is identified by a set of spacetime element faces that precisely cover the trajectory of the contact discontinuity, rather than by mesh

refinement. We used the new capability to mesh with inclined tent poles to align the spacetime mesh with the motion of the discontinuity. Since the singular surface aligns with element boundaries in the spacetime mesh, the SDG formulation weakly enforces the correct set of jump conditions at the contact discontinuity. Notably, the contact discontinuity is captured as a true jump in the SDG solution (Fig. 10) — without mesh refinement and with no extra stabilization active beyond the inherent stability of the SDG formulation. Thus, we obtain a more accurate resolution of the contact discontinuity at substantially lower cost than would be required if a capturing technique were applied to this feature.

We used *a priori* knowledge of the motion of the contact discontinuity to construct the spacetime mesh in this proof-of-concept study. However, the motion of discontinuities is often solution-dependent, so a more robust approach is generally required. We are currently testing a more robust approach, in which the inclination of the tent pole in each patch is iteratively adjusted to satisfy the zero-mass-flux condition that governs the motion of the contact discontinuity. We plan to use this method and related techniques in continuing work to track moving interfaces in materials microstructures and to follow crack growth along solution-dependent paths.

6 Conclusions

We have demonstrated three distinct approaches to discretizing evolving discontinuous solution features using SDG solution methods. The first option, h-adaptive procedures for capturing discontinuities, mitigates the cost of the necessary grid refinement by adapting the mesh simultaneously in space and time, thereby avoiding the expense imposed by a global uniform time step dictated by the smallest elements in a spatially adapted discretization. The second approach, the incorporation of a cohesive damage model in the SDG framework, is facilitated by the discontinuous SDG solution spaces and the treatment of jump conditions that naturally accommodate the enforcement of the traction–separation law. We included an adaptive error indicator that directly ensures the accuracy of the finite element approximation of the TSL. The high resolution obtained with this model led to the discovery that the material velocity exhibits a strong singular response for running cracks with the same $r^{-\frac{1}{2}}$ structure observed in classical elastodynamic fracture. We also presented a preliminary study of how the use of inclined tent poles and a more robust set of adaptive meshing procedures can be used to track moving discontinuities. This approach uses the discontinuous nature of SDG solution spaces to eliminate the need for strong mesh refinement near the discontinuity, to enforce the correct jump conditions dictated by the governing balance laws, and to render a sharp resolution of the discontinuity. It's worth noting that our software framework allows us to combine more than one of these

techniques to address different classes of evolving discontinuities in a single computation.

Execution times for the numerical examples reported here ranged from several hours to several days of serial processing on desk-top processors; the longest were for the highly refined multi-scale model that resolves the singular velocity field at a moving cohesive crack tip. These relatively long execution times are for a non-optimized C++ research code and are not representative of the true cost of the SDG method. Recent improvements to our serial code yielded substantial speed-ups in the range of $30 - 40$ times faster. The optimizations included use of the Blitz++ library [47] to mitigate some of the inefficiencies of array storage in C++, more effective use of standard math libraries, use of more aggressive compiler optimizations, and straightforward eliminations of redundant calculations. Additional serial speed-ups are expected as we continue to optimize the code. Parallel implementation is another promising direction to improve performance. We obtain high efficiency and nearly linear speed-ups in the number of parallel processors for non-adaptive computations. The main challenge for an effective adaptive parallel implementation is maintaining load balance in the face of intense and evolving adaptive refinement. We believe that the local nature of our patch-by-patch solution scheme will prove useful in on-going research that addresses load balancing in parallel–adaptive SDG computations.

We continue to develop the underlying SDG formulation and to extend the technologies for discontinuity capturing reported in this work. One area of continuing effort involves improved methods for SDG shock capturing to control overshoot in nonlinear conservation laws. We seek methods that are suitable for spacetime implementation, that do not interfere with the element-wise conservation property, that preserve the compact form of the SDG stencil, and that introduce as little artificial viscosity as possible. SDG variants of sub-cell shock capturing strategies [39] appear particularly promising at this time. We continue to develop the SDG technology for tracking moving interfaces and discontinuities. One such improvement involves the use of interface kinetics models to track solution-dependent motions of interfaces and discontinuities. Another application is the use of the interface tracking technology to track cohesive crack growth along arbitrary, solution-dependent paths (vs. the predetermined paths in our current studies). Extending our h-adaptive technology to hp enrichments is another promising direction for development.

Acknowledgements

The authors gratefully acknowledge the contributions of Jeff Erickson, Michael Garland, Shripad Thite and Yuan Zhou to this work. Support from the Center for Process Simulation and Design (CPSD) and the Center for Simulation of Advanced Rockets (CSAR) at the University of Illinois at Urbana-Champaign is gratefully acknowledged. The U.S. National Science Foundation supports

research in CPSD *via* grant NSF DMR 01-21695. The CSAR research program is supported by the U.S. Department of Energy through the University of California under subcontract B341494.

References

1. Abedi R, Chung S-H, Erickson J, Fan Y, Garland M, Guoy D, Haber R, Sullivan J, Thite S, and Zhou Y (2004) Proc 20th Ann ACM Symp on Comp Geometry 300–309
2. Abedi R, Haber RB, Petracovici B (2006) Comp Methods Appl Mechs Engrg 195: 3247–3273.
3. Abedi R, Haber RB, Thite S, Erickson J (2006) Rev Europ Mécanique Num 15:619–642
4. Babuška I, Melenk I (1997) Int J Num Methods Engrg 40:727–758
5. Barenblatt GI (1962) Adv Appl Mechs 7:55–129.
6. Belytschko T, and Black T (1999) Int J Num Methods Engrg 45:601–620
7. Bishop RL, Goldberg SI (1980) Tensor analysis on manifolds. Prentice Hall (reprinted by Dover), Englewood Cliffs, New Jersey
8. Brezzi F, Bristeau M-O, Franca L, Mallet M, Rogé G (1992) Comp Methods Appl Mechs Engrg 96:117–129
9. Brooks AN, Hughes TJR (1982) Comp Methods Appl Mechs Engrg 32:199–259
10. Camacho GT, Ortiz M (1996) Int J Solids Structs 33:2899–2938
11. Cockburn B, Gremaud P-A (1996) SIAM J Num Analysis 33:522–554
12. Cockburn B, Karniadakis GE, Shu CW (eds.) (2000) Discontinuous Galerkin methods -theory, computation and applications, Lecture Notes in Computational Science and Engineering 11, Springer-Verlag, New York
13. Cornec A, Scheider I, Schwalbe K-H (2003) Engrg Fracture Mechs 70:1963–1987
14. de Borst R, Remmers JJC, Needleman A (2006) Engrg Fracture Mechs 73:160–177
15. Dugdale DS (1960) J Mechs Phys Solids 8:100–104
16. Erickson J, Guoy D, Sullivan J, Üngör A (2002) Proc 11th Int Meshing Roundtable (Sandia National Laboratories) 391–402
17. Fedkiw RP, Aslam T, Merriman B, Osher S (1999) J. Comp. Physics 152:457–492
18. Fleming WH (1964) Functions of several variables, Addison-Wesley, Reading,Massachusetts
19. Freund LB (1990) Dynamic Fracture Mechanics, Cambridge University Press, Cambridge, UK
20. Geubelle PH, Baylor J, (1998) Composites 29B:589–602
21. Huang H, Costanzo F (2004) Int J Fracture 127:119–146
22. Hughes TJR, Liu WK, Zimmermann TK (1981) Comp Methods Appl Mechs Engnrg 29:329–349
23. Hughes TJR, Franca LP, Hulbert GM (1989) Comp Methods Appl Mechs Engnrg 73:173–189
24. Hughes TJR, Mallet M (1986) Comp Methods Appl Mechs Engnrg 58:305–328
25. Jaffré J, Johnson C, Szepessy A (1995) Math Models Methods Appl Sci 5:367–386

26. Johnson C, Szepessy A (1987) Math Comp 49:427–444
27. Karamanos G-S, Karniadakis GE (2000) J Comp Physics 163:22–50
28. Krysl P, Belytschko T (1999) Int J Num Methods Engrg 44:767–800
29. Lowrie RB, Roe PL, van Leer B (1998) in Barriers and Challenges in Computational Fluid Dynamics, ICASE/LaRC Interdisciplinary Series in Science and Engineering, Natl Inst Aerospace 6:79–98
30. Luo H, Baum JD, Löhner R (2004) J Comp Physics 194:304–328
31. Mergheim J, Kuhl E, Steinmann P (2005) Int J Num Methods Engrg 63:276–289
32. Moës N, Belytschko T (2002) Eng Fracture Mechs 69:813–833
33. Moës N, Dolbow J, Belytschko T (1999) Int J Num Methods Engrg 46:131–150
34. Needleman A (1987) J Appl Mechs 54:525–532
35. Ortiz M, Pandolfi A (1999) Int J Num Methods Engrg 44:1267–1282
36. Osher S, Sethian JA (1988) J Comp Physics 79:12–49
37. Palaniappan J, Haber RB, Jerrard RL (2004) Comp Methods Appl Mechs Engnrg 193:3607–3631
38. Pandolfi A, Ortiz M (2002) Engnrg Computers 18:148–159
39. Persson P-P, Peraire J (2006) 44th AIAA Aerospace Sci Mtg Exhibit, AIAA, AIAA–2006–0012
40. Richter GR (1994) Appl Num Math 16:65–80
41. Schrefler BA, Secchi S, Simoni L (2006) Comp Methods Appl Mechs Engrg 195:444–461
42. Shakib F, Hughes TJR, Johan A (1991) Comp Methods Appl Mechs Engrg 89:141–219
43. Spivak M (1965) Calculus on manifolds, W. A. Benjamin, New York
44. Tadmor E (1989) SIAM J Num Analysis 26:30–44
45. Tvergaard V, Hutchinson JW (1992) J Mechs Phys Solids 40:1377–1397
46. Üngör A, Sheffer A (2002) Int J Foundations Comp Sci 13: 201–221
47. Veldhuizen TL (1998) Arrays in Blitz++. In: Caromel D, Oldehoeft RR, Tholburn M (eds) Proc 2nd Int Symp on Scientific Computing in Object-Oriented Parallel Environments (ISCOPE'98), Lecture Notes in Computer Science 1505, Springer-Verlag, London UK
48. Wells GN, Sluys LJ (2001) Int J Num Methods Engrg 50:2667–2682
49. Xu X-P, Needleman A (1994) J Mechs Phys Solids 42:1397–1434
50. Yin L, Acharya A, Sobh N, Haber RB, and Tortorelli DA (2000) A spacetime discontinuous Galerkin method for elastodynamics analysis. In: Cockburn B, Karniadakis G, Shu CW (eds.) Discontinuous Galerkin methods: Theory, computation and applications, Lecture Notes in Computational Science and Engineering 11, Springer-Verlag, Berlin
51. Zhou F, Molinari J-F, Shioya T (2005) Engrg Fracture Mechs 72:1383–1410
52. Zhou Y, Garland M, Haber RB (2004) Pixel-exact rendering of spacetime finite element solutions, In: Proc IEEE Visualization 2004, IEEE 425–432
53. Zi G, Belytschko T (2003) Int J Num Methods Engrg 57:2221–2240.

Analysis of a finite element formulation for modelling phase separation

Garth N. Wells[1] and Krishna Garikipati[2]

[1] Delft University of Technology g.n.wells@tudelft.nl
[2] University of Michigan krishna@umich.edu

Summary. The Cahn-Hilliard equation is of importance in materials science and a range of other fields. It represents a diffuse interface model for simulating the evolution of phase separation in solids and fluids, and is a nonlinear fourth-order parabolic equation, which makes its numerical solution particularly challenging. To this end, a finite element formulation has been developed which can solve the Cahn-Hilliard equation in its primal form using C^0 basis functions. Here, analysis of a fully discrete version of this method is presented in the form of *a priori* uniqueness, stability and error analysis.

Key words: Cahn-Hilliard equation, discontinuous Galerkin method, phase separation.

1 Introduction

The Cahn-Hilliard [1] equation models the separation of phases in binary mixtures. It is particularly relevant in material science, where it describes microstructure evolution in alloys. As a diffuse interface model, it represents the boundaries between pure phases as a small region across which the relative concentration varies rapidly. This is an advantage in that discontinuities at phase boundaries do not have to be modelled explicitly, but comes at the cost of needing to resolve gradients at phase boundaries well and a high degree of mathematical complexity.

The numerical solution of the Cahn-Hilliard equation is particularly challenging on a number of fronts. Foremost in the difficulties is the presence of fourth-order spatial derivatives. A fourth-order term is necessary to balance the presence of a chemical potential in the Cahn-Hilliard equation which is derived from a non-convex free energy function. The non-convex nature of the chemical free energy further complicates both the numerical solution strategy and the analysis of numerical methods. Extensive research has been

Alain Combescure et al. (eds.), IUTAM Symposium on Discretization Methods for Evolving Discontinuities, 89–102.

performed over the past 20 years into the mathematical analysis of finite element formulations for the Cahn-Hilliard equation. Initial efforts focused upon C^1 [2, 3] and non-conforming methods [4]. The majority of later efforts focused upon a mixed finite element formulation based on an operator split (see Refs. [5, 6, 7, 8] for a selection of works).

We present here analysis of the finite element formulation for the Cahn-Hilliard equation presented by Wells et al. [9]. The formulation allows the solution of the Cahn-Hilliard equation using standard C^0 basis functions in the primal form, with continuity of the normal derivative across element boundaries enforced weakly. In Wells et al. [9], stability and convergence in an energy-like norm was demonstrated for the time-continuous case. Here, a fully discrete formulation is analysed and an error estimate in the L^2 norm is presented.

The rest of this work is organised as follows: the Cahn-Hilliard equation is presented in its strong form, after which the semi-discrete Galerkin formulation is presented and some key results are summarised. The fully discrete formulation is then presented, followed by *a priori* analysis of uniqueness, stability and error in the L^2 norm. A numerical example is presented, simulating phase separation in a uniform mixture which is randomly-perturbed. Following the numerical example, conclusions are drawn.

2 Cahn-Hilliard equation

Consider a binary mixture and let the concentration of one of its constituents, say A, be denoted by c satisfying $0 < c < 1$. The concentration of the other constituent, B, is $1 - c$. Pure phases are obtained for $c = 0$ and $c = 1$. Let the mixture occupy an open, simply connected region in space, $\Omega \subset \mathbb{R}^d$, where $d = 1$, 2 or 3. The boundary of Ω is supposed to be sufficiently smooth, and is denoted by $\Gamma = \partial\Omega$, with outward unit normal $\boldsymbol{n}$. In strong form we have the following problem: find $c : \overline{\Omega} \times [0, T] \to \mathbb{R}$ such that

$$
\begin{aligned}
& c_{,t} = \nabla \cdot \left(M \nabla \left(\mu - \lambda \nabla^2 c \right) \right) && \text{in } \Omega \times (0, T), && (1)\\
& \nabla c \cdot \boldsymbol{n} = 0 && \text{on } \Gamma \times (0, T), && (2)\\
& \nabla \left(\mu - \lambda \nabla^2 c \right) \cdot \boldsymbol{n} = s && \text{on } \Gamma \times (0, T), && (3)\\
& c(\boldsymbol{x}, 0) = c_0(\boldsymbol{x}) && \text{in } \Omega, && (4)
\end{aligned}
$$

where $M \geq 0$ is the mobility, μ is the chemical potential, $\lambda > 0$ is a constant related to the interfacial energy and $c_0(\boldsymbol{x})$ are the initial conditions. The boundary conditions of zero normal derivative for the concentration and zero mass flux on the entire boundary are typical for the Cahn-Hilliard equation. The Cahn-Hilliard equation is a fourth-order diffusion equation, and the presence of fourth-order spatial derivatives is due to the introduction of an expression for the surface free-energy Ψ^s,

$$
\Psi^s - \frac{\lambda}{2} \nabla c \cdot \nabla c. \tag{5}
$$

The mobility M can be either constant or concentration-dependent. If made concentration-dependent, a common choice is

$$M = c\,(1-c)\,, \tag{6}$$

which largely restricts diffusion processes to phase interface zones. The chemical potential in the case of phase separation problems comes from a non-convex chemical free-energy Ψ^c. It is the non-convex nature of the free-energy which drives phase separation. A typical chemical free-energy of a solution is given by [1]

$$\Psi^c = NkT\,(c\ln c + (1-c)\ln(1-c)) + N\omega c\,(1-c)\,, \tag{7}$$

where N is the number of molecules per unit volume, k is Boltzmann's constant, T is the absolute temperature and ω is a parameter related to the mixing enthalpy that determines the shape of Ψ^c. For $\omega > 2kT$, the chemical free energy is non-convex, with two wells close to $c = 0$ and $c = 1$ which drives phase segregation into the two binodal points. The potential μ is given by the functional derivative of the chemical free-energy with respect to the concentration, $D_c\Psi^c$. Another possibility is the use of a quartic polynomial for the chemical free energy,

$$\Psi^c = \frac{1}{4}c^2\,(1-c)^2\,, \tag{8}$$

which has the advantage of being continuous on the real line. This simplifies the analysis of formulations and leads to more robust numerical procedures. It does however allow for solutions outside of the range $[0, 1]$. The term λ governs the magnitude of the free-energy related the interfaces in the presence of a given concentration gradient.

3 Semi-discrete Galerkin formulation

3.1 Definitions

Consider a partition of Ω into n_{el} polygonal open sets, Ω_e, each with boundary $\Gamma_e = \partial\Omega_e$:

$$\Omega = \overline{\bigcup_{e=1}^{n_{\text{el}}} \Omega_e}, \quad \text{such that } \bigcap_{e=1}^{n_{\text{el}}} \Omega_e = \emptyset. \tag{9}$$

It is assumed that Ω is a polygon and hence can be partitioned exactly. The union of inter-element boundaries and the boundary Γ is denoted by

$$\widetilde{\Gamma} = \bigcup_{e=1}^{n_{\text{el}}} \Gamma_e, \tag{10}$$

where $\Gamma_e = \partial\Omega_e$, and the union of element interiors is denoted by

$$\widetilde{\Omega} = \bigcup_{e=1}^{n_{\text{el}}} \Omega_e. \tag{11}$$

The jump operator for a vector is denoted by

$$[\![\boldsymbol{a}]\!] = \boldsymbol{a}_1 \cdot \boldsymbol{n}_1 + \boldsymbol{a}_2 \cdot \boldsymbol{n}_2, \quad \text{on } \widetilde{\Gamma} \backslash \Gamma, \tag{12}$$

$$[\![\boldsymbol{a}]\!] = \boldsymbol{a} \cdot \boldsymbol{n} \quad \text{on } \Gamma, \tag{13}$$

where the subscripts refer to the face of the element on either side of each inter-element boundary, and $\boldsymbol{n}$ is the unit outward normal to an element boundary. The average operator is denoted by

$$\langle a \rangle = \frac{1}{2}(a_1 + a_2) \quad \widetilde{\Gamma} \backslash \Gamma, \tag{14}$$

$$\langle a \rangle = a \quad \text{on } \Gamma, \tag{15}$$

where again the subscripts refer to the face of the element on either side of each inter-element boundary.

3.2 Semi-discrete formulation

Classical Galerkin methods for the Cahn-Hilliard equation with the considered boundary conditions seek approximate solutions in a subspace of $H^2_E(\Omega)$, which is defined as

$$H^2_E(\Omega) = \left\{ c^h \middle| \, c^h \in H^2(\Omega), \ \nabla c^h \cdot \boldsymbol{n} \text{ on } \Gamma \right\}. \tag{16}$$

The space satisfies the considered Dirichlet boundary condition by construction. However, in a finite element context, such functions are difficult to construct. Here, a Galerkin formulation for the Cahn-Hilliard equation is examined which looks for solutions in a subspace of $H^1(\Omega)$, thereby allowing the use of standard C^0 Lagrange shape functions. Consider therefore the function space

$$W^h = \left\{ c^h \middle| \, c^h \in H^1(\Omega), \ c^h \in P^k(\Omega_e) \, \forall e \right\}, \tag{17}$$

where $P^k(\Omega_e)$ is the space of the standard polynomial finite element shape functions on element Ω_e and k is the polynomial order. Note that $c^h \in H^2\left(\widetilde{\Omega}\right)$. A finite element problem for the Cahn-Hilliard equation then involves [9]: find $c^h(t) \in W^h$, $t \in [0, T]$ such that

$$\left(w^h, \dot{c}^h\right)_{\Omega} + a\left(w^h, c^h\right) = 0 \quad \forall \ w^h \in W^h, \tag{18}$$

where

$$\begin{aligned} a\left(w^h, c^h\right)_{\Omega} = & \left(\nabla w^h, M^h \nabla \mu^h\right)_{\Omega} + \left(\nabla^2 w^h, M^h \lambda \nabla^2 c^h\right)_{\widetilde{\Omega}} \\ & + \left(\nabla w^h, \left(\nabla M^h\right) \lambda \nabla^2 c^h\right)_{\widetilde{\Omega}} - \left([\![\nabla w^h]\!], \langle M^h \lambda \nabla^2 c^h \rangle\right)_{\widetilde{\Gamma}} \\ & - \left(\langle M^h \lambda \nabla^2 w^h \rangle, [\![\nabla c^h]\!]\right)_{\widetilde{\Gamma}} + \frac{\alpha}{h}\left([\![\nabla w^h]\!], M^h \lambda [\![\nabla c^h]\!]\right)_{\widetilde{\Gamma}}, \end{aligned} \tag{19}$$

and α is a dimensionless penalty term and h is a measure of the element size. The notation

$$(a,b)_X = \int_X ab\, dX \tag{20}$$

for inner products has been adopted. In effect, the formulation imposes continuity of (a) the normal derivative of the concentration, (b) the normal flux and (c) the Laplacian of the concentration across element boundaries in a weak sense. For simplicity of notation, it has been assumed that α and h are constant for all elements.

The formulation can be shown to be consistent with the Cahn-Hilliard equation, and is stable if the penalty term is chosen to be sufficiently large [9]. How large it must be is dependent on constants in various inequalities which are related to the order of elements and the element geometry. Stability estimates and *a priori* error estimates in an energy-like norm for the semi-discrete problem can be found in Wells et al. [9]. The focus in the rest of this work is on estimates for the fully discrete problem.

4 Fully discrete formulation

We consider now the stability of a time-discrete problem whose numerical scheme is parametrised by $\theta \in [0,1]$. The time continuous problem in equation (18) is replaced by a sequence of discrete steps at $t_1, t_2, \cdots, t_n, t_{n+1}$, where $\Delta t = t_{n+1} - t_n$. The problem at t_{n+1} then becomes: find $c^{h,n+1} \in W^h$ such that

$$\left(w^h, \frac{c^{h,n+1} - c^{h,n}}{\Delta t}\right) + a\left(w^h, c^{h,n+\theta}\right) = 0 \quad \forall\, c^h \in W^h, \tag{21}$$

where

$$c^{h,n+\theta} = (1-\theta)\, c^{h,n} + \theta\, c^{h,n+1}. \tag{22}$$

As usual, $\theta = 0$ leads to the forward Euler scheme, $\theta = 1$ leads to the backward Euler scheme and $\theta = 1/2$ leads to the Crank-Nicolson method.

In the following analysis, we restrict ourselves to the case of constant mobility ($M = 1$) and a continuously differentiable chemical potential with the property

$$\frac{d\mu}{dc} \geq -b_0, \tag{23}$$

where $b_0 \geq 0$. This condition holds for all commonly adopted chemical potentials.

4.1 Uniqueness

Consider the function $\beta = c_1^{h,n+1} - c_2^{h,n+1}$, where $c_1^{h,n+1}$ and $c_2^{h,n+1}$ are solutions to the fully discrete problem. Inserting β into the right-hand slot of

the forms in equation (21), and noting the $w^h = \beta$ is an admissible weighting function,

$$\|\beta\|_{\Omega}^2 + \Delta t\theta\lambda \left\|\nabla^2\beta\right\|_{\widetilde{\Omega}}^2 - 2\Delta t\theta \left(\left[\!\left[\nabla\beta\right]\!\right], \left\langle\nabla^2\beta\right\rangle\right)_{\widetilde{\Gamma}} + \Delta t\theta\frac{\alpha\lambda}{h} \left\|\left[\!\left[\nabla\beta\right]\!\right]\right\|_{\widetilde{\Gamma}}^2 = -\Delta t \left(\nabla\beta, \nabla\left(\mu\left(c_1^{h,n+\theta}\right) - \mu\left(c_2^{h,n+\theta}\right)\right)\right)_{\Omega}, \quad (24)$$

where $\|a\|_X$ indicates the L^2 norm on the domain X. From integration by parts,

$$\left|\left(\nabla\beta, \nabla\left(\mu\left(c_1^{h,n+\theta}\right) - \mu\left(c_2^{h,n+\theta}\right)\right)\right)_{\Omega}\right| \leq \left|\left(\nabla^2\beta, \mu\left(c_1^{h,n+\theta}\right) - \mu\left({}^{h,n+\theta}c_2\right)\right)_{\widetilde{\Omega}}\right| + \left|\left(\left[\!\left[\nabla\beta\right]\!\right], \mu\left(c_1^{h,n+\theta}\right) - \mu\left(c_2^{h,n+\theta}\right)\right)_{\widetilde{\Gamma}}\right|, \quad (25)$$

together with Lipschitz continuity of the chemical potential,

$$\begin{aligned}\left|\mu\left(c_1^{h,n+\theta}\right) - \mu\left(c_2^{h,n+\theta}\right)\right| &\leq L\left|c_1^{h,n+\theta} - c_2^{h,n+\theta}\right| \\ &= L\theta\left|\beta\right|,\end{aligned} \quad (26)$$

where $L > 0$, and the Cauchy-Schwartz inequality, it follows that:

$$\left|\left(\nabla\beta, \nabla\left(\mu\left(c_1^{h,n+\theta}\right) - \mu\left(c_2^{h,n+\theta}\right)\right)\right)_{\Omega}\right| \leq \theta L \left\|\nabla^2\beta\right\|_{\widetilde{\Omega}} \|\beta\|_{\Omega} + \theta L \left\|\left[\!\left[\nabla\beta\right]\!\right]\right\|_{\widetilde{\Gamma}} \|\beta\|_{\widetilde{\Gamma}}. \quad (27)$$

Application of Young's inequality then leads to:

$$\left|\left(\nabla\beta, \nabla\left(\mu\left(c_1^{h,n+\theta}\right) - \mu\left(c_2^{h,n+\theta}\right)\right)\right)_{\Omega}\right| \leq \frac{\theta L}{2\epsilon_0}\left\|\nabla^2\beta\right\|_{\widetilde{\Omega}}^2 + \frac{\theta L\epsilon_0}{2}\|\beta\|_{\Omega} + \frac{\theta L}{2\epsilon_1}\left\|\left[\!\left[\nabla\beta\right]\!\right]\right\|_{\widetilde{\Gamma}}^2 + \frac{\theta L\epsilon_1}{2}\|\beta\|_{\widetilde{\Gamma}}^2, \quad (28)$$

where $\epsilon_0, \epsilon_1 > 0$. Turning attention now to the first inter-element term in equation (24),

$$2\left|\left(\left[\!\left[\nabla\beta\right]\!\right], \langle\nabla\beta\rangle\right)_{\widetilde{\Gamma}}\right| \leq \epsilon_2 C\left\|\nabla^2\beta\right\|_{\widetilde{\Omega}}^2 + \frac{1}{h\epsilon_2}\left\|\left[\!\left[\nabla\beta\right]\!\right]\right\|_{\widetilde{\Gamma}}^2. \quad (29)$$

For the derivation of this expression, we refer the reader to Wells et al. [9]. Also from Wells et al. [9], there exists $C > 0$ such that

$$\|\beta\|_{\widetilde{\Gamma}}^2 \leq \frac{C}{h}\|\beta\|_{\Omega}^2. \quad (30)$$

Using these results in equation (24), it follows that

$$\|\beta\|^2_{\Omega} + \Delta t\theta\lambda \left\|\nabla^2\beta\right\|^2_{\widetilde{\Omega}} + \Delta t\theta \frac{\alpha\lambda}{h} \|[\![\nabla\beta]\!]\|^2_{\widetilde{\Gamma}}$$
$$\leq \frac{\Delta t\theta L}{2\epsilon_0} \left\|\nabla^2\beta\right\|^2_{\widetilde{\Omega}} + \frac{\Delta t\theta L\epsilon_0}{2} \|\beta\|^2_{\Omega} + \frac{\Delta t\theta L}{2\epsilon_1} \|[\![\nabla\beta]\!]\|^2_{\widetilde{\Gamma}}$$
$$+ \frac{\Delta t\theta L\epsilon_1 C}{2h} \|\beta\|^2_{\Omega} + \Delta t\theta\lambda\epsilon_2 C \left\|\nabla^2\beta\right\|^2_{\widetilde{\Omega}} + \frac{\Delta t\theta\lambda}{h\epsilon_2} \|[\![\nabla\beta]\!]\|^2_{\widetilde{\Gamma}}. \quad (31)$$

Grouping related terms together,

$$\|\beta\|^2_{\Omega} + \left(\Delta t\theta\lambda - \frac{\Delta t\theta L}{2\epsilon_0} - \lambda\Delta t\theta C\epsilon_2\right) \left\|\nabla^2\beta\right\|^2_{\widetilde{\Omega}}$$
$$+ \left(\frac{\Delta t\theta\alpha\lambda}{h} - \frac{\Delta t\theta L}{2\epsilon_1} - \frac{\Delta t\theta\lambda}{h\epsilon_2}\right) \|[\![\nabla\beta]\!]\|^2_{\widetilde{\Gamma}}$$
$$\leq \left(\frac{\Delta t\theta L\epsilon_0}{2} + \frac{\Delta t\theta L\epsilon_1 C}{2h}\right) \|\beta\|^2_{\Omega}. \quad (32)$$

The goal now is to select ϵ_i such that the RHS is greater than $\|\beta\|^2_{\Omega}$ and all terms on the LHS are positive. Setting $\epsilon_0 = 1/\Delta t\theta L$, $\epsilon_1 = h/\Delta t\theta LC$ and $\epsilon_2 = \epsilon_2'/C$

$$\|\beta\|^2_{\Omega} + \theta \left(\Delta t\lambda - \frac{\Delta t^2\theta L^2}{2} - \lambda\Delta t\epsilon_2'\right) \left\|\nabla^2\beta\right\|^2_{\widetilde{\Omega}}$$
$$+ \theta \left(\frac{\Delta t\alpha\lambda}{h} - \frac{\Delta t^2\theta L^2 C}{2h} - \frac{\Delta t\lambda C}{h\epsilon_2'}\right) \|[\![\nabla\beta]\!]\|^2_{\widetilde{\Gamma}} \leq \|\beta\|^2_{\Omega}. \quad (33)$$

To demonstrate uniqueness, Δt and α must be chosen such that all terms on the LHS are greater than or equal to zero. Consider therefore the restriction on the time step

$$\Delta t < \frac{2\lambda\,(1-\epsilon_2')}{\theta L^2}, \quad (34)$$

where $0 < \epsilon_2' < 1$. Subject to this time step restriction and if

$$\alpha > (1-\epsilon_2')\,C + \frac{C}{\epsilon_2'} \quad (35)$$

it follows that

$$\|\beta\|^2_{\Omega} + a_1 \left\|\nabla^2\beta\right\|^2_{\widetilde{\Omega}} + a_2 \|[\![\nabla\beta]\!]\|^2_{\widetilde{\Gamma}} \leq \|\beta\|^2_{\Omega}, \quad (36)$$

where $a_1, a_2 > 0$, which can hold only if $\beta = 0$. Therefore, under the time step restriction, and for a sufficiently large penalty, the fully discrete scheme has a unique solution. The analysis indicates that a larger penalty allows for a larger time step since a large α allows for a small ϵ_2'. Note that the time step restriction is not dependent on the element size h. It is a function of model and time stepping parameters only, whereas α is dependent only on element shape parameters and not model parameters.

The time step restriction for uniqueness is due to the non-convex nature of Ψ^c. It is possible that a more subtle analysis of the problem may lead to a tighter bound for the maximum allowable time step.

4.2 Stability

Stability of the time discrete formulation is now considered by setting $w^h = c^{h,n+\theta}$ in equation (21),

$$\left(c^{h,n+\theta}, \frac{c^{h,n+1} - c^{h,n+1}}{\Delta t}\right)_\Omega + a\left(c^{h,n+\theta}, c^{h,n+\theta}\right) = 0. \tag{37}$$

Noting that

$$c^{h,n+\theta} = \left(\theta - \frac{1}{2}\right)\left(c^{h,n+1} - c^{h,n}\right) + \frac{c^{h,n+1} + c^{h,n}}{2}, \tag{38}$$

equation (37) can be expressed as:

$$\left(\theta - \frac{1}{2}\right)\left\|c^{h,n+1} - c^{h,n}\right\|_\Omega^2 + \frac{\left\|c^{h,n+1}\right\|_\Omega^2 - \left\|c^{h,n}\right\|_\Omega^2}{2\Delta t} + a\left(c^{h,n+\theta}, c^{h,n+\theta}\right) = 0. \tag{39}$$

For the case $\theta \in [1/2, 1]$,

$$\frac{\left\|c^{h,n+1}\right\|_\Omega^2 - \left\|c^{h,n}\right\|_\Omega^2}{2\Delta t} + a\left(c^{h,n+\theta}, c^{h,n+\theta}\right) \leq 0. \tag{40}$$

For standard parabolic differential equations, $a\left(c^{h,n+\theta}, c^{h,n+\theta}\right) \geq 0$, leading trivially to $\left\|c^{h,n+1}\right\|^2 \leq \left\|c^{h,n}\right\|^2$ which implies stability. However, this is not the case for the Cahn-Hilliard equation as the term $a\left(c^{h,n+\theta}, c^{h,n+\theta}\right)$ may be negative. Demonstrating stability requires a more subtle approach.

Expanding the term $a\left(c^{h,n+\theta}, c^{h,n+\theta}\right)$ and rearranging equation (40),

$$\begin{aligned}
\frac{1}{2\Delta t}\left\|c^{h,n+1}\right\|_\Omega^2 + \lambda\left\|\nabla^2 c^{h,n+\theta}\right\|_{\widetilde{\Omega}}^2 + \frac{\alpha\lambda}{h}\left\|\left[\!\left[\nabla c^{h,n+\theta}\right]\!\right]\right\|_{\widetilde{\Gamma}}^2 \\
\leq \frac{b_0}{2\epsilon_0}\left\|\nabla^2 c^{h,n+\theta}\right\|_{\widetilde{\Omega}}^2 + \frac{b_0}{2\epsilon_1}\left\|\left[\!\left[\nabla c^{h,n+\theta}\right]\!\right]\right\|_{\widetilde{\Gamma}}^2 + \frac{b_0\epsilon_0}{2}\left\|c^{h,n+\theta}\right\|_\Omega^2 \\
+ \frac{b_0 C\epsilon_1}{2h}\left\|c^{h,n+\theta}\right\|_\Omega^2 + \lambda\epsilon_2 C\left\|\nabla^2 c^{h,n+\theta}\right\|_{\widetilde{\Omega}}^2 \\
+ \frac{\lambda}{h\epsilon_2}\left\|\left[\!\left[\nabla c^{h,n+\theta}\right]\!\right]\right\|_{\widetilde{\Gamma}}^2 + \frac{1}{2\Delta t}\left\|c^{h,n}\right\|_\Omega^2,
\end{aligned} \tag{41}$$

where $\epsilon_i > 0$. Setting now $\epsilon_0 = b_0/2\lambda\epsilon_0'$, $\epsilon_1 = b_0 h/2C\lambda\epsilon_1'$, and $\epsilon_2 = \epsilon_2'/C$, it follows that

$$\begin{aligned}
\frac{1}{2\Delta t}\left\|c^{h,n+1}\right\|_\Omega^2 + \lambda\left(1 - \epsilon_0' - \epsilon_1'\right)\left\|\nabla^2 c^{h,n+\theta}\right\|_{\widetilde{\Omega}}^2 \\
+ \frac{\lambda}{h}\left(\alpha - C\epsilon_1' - \frac{C}{\epsilon_2'}\right)\left\|\left[\!\left[\nabla c^{h,n+\theta}\right]\!\right]\right\|_{\widetilde{\Gamma}}^2 \\
\leq \left(\frac{b^2}{4\lambda\epsilon_0'} + \frac{b^2}{4\lambda\epsilon_1'}\right)\left\|c^{h,n+\theta}\right\|_\Omega^2 + \frac{1}{2\Delta t}\left\|c^{h,n}\right\|_\Omega^2.
\end{aligned} \tag{42}$$

Setting $\epsilon_0' = \epsilon_1' = 1 - \epsilon_2'$, the term $\lambda\left(1 - \epsilon_0' - \epsilon_1'\right)\left\|\nabla^2 c^{h,n+\theta}\right\|_{\widetilde{\Omega}}^2$ vanishes, and if

$$\alpha > \left(1 - \epsilon_2'\right) C + \frac{C}{\epsilon_2'}, \tag{43}$$

then $(\lambda/h)\left(\alpha - C\epsilon_1' - C/\epsilon_2'\right)\left\|\left[\!\left[\nabla c^{h,n+\theta}\right]\!\right]\right\|_{\widetilde{\Gamma}}^2 > 0$. It follows then from equation (42) that

$$\left\|c^{h,n+1}\right\|_{\Omega}^2 \leq \frac{\Delta t b^2}{\lambda\left(1 - \epsilon_2'\right)}\left\|c^{h,n+\theta}\right\|_{\Omega}^2 + \left\|c^{h,n}\right\|_{\Omega}^2. \tag{44}$$

For the case $\theta = 1$, this implies that if

$$\Delta t < \frac{\lambda\left(1 - \epsilon_2'\right)}{b^2} \tag{45}$$

then the method is stable as the solution at time step $n+1$ is bounded in terms of the solution at time step n, although it does not imply $\left\|c^{h,n+1}\right\|_{\Omega} \leq \left\|c^{h,n}\right\|_{\Omega}$. For other cases,

$$\left\|c^{h,n+\theta}\right\|_{\Omega}^2 \leq 2\left(\left\|(1-\theta)\, c^{h,n}\right\|_{\Omega}^2 + \left\|\theta c^{h,n+1}\right\|_{\Omega}^2\right). \tag{46}$$

Therefore, for $\theta = 1/2$ the critical time step is the same as for the $\theta = 1$ case. Stability is assured if the critical time step is met since when summing over all time steps, the solution remains bounded in terms of the initial conditions.

Due to the presence of second-order derivatives in the weak form, for the case $\theta \in [0, 1/2)$ the allowable time step $\Delta t \propto h^4$. This makes the usefulness of such schemes extremely limited. Hence, the analysis of such schemes is not pursued here.

4.3 Accuracy

Consider the elliptic projection $P^h u : H_E^2(\Omega) \to W^h$ defined by: given $u \in H_E^2(\Omega)$, find $P^h u \in W^h$ such that

$$\begin{aligned}\left(\nabla^2 w^h, \nabla^2\left(P^h u - u\right)\right)_{\widetilde{\Omega}} - \left(\left[\!\left[\nabla w^h\right]\!\right], \left\langle\nabla^2\left(P^h u - u\right)\right\rangle\right)_{\widetilde{\Gamma}} \\ - \left(\left\langle\nabla^2 w^h\right\rangle, \left[\!\left[\nabla^2\left(P^h u - u\right)\right]\!\right]\right)_{\widetilde{\Gamma}} \\ + \frac{\alpha}{h}\left(\left[\!\left[\nabla w^h\right]\!\right], \left[\!\left[\nabla P^h u\right]\!\right]\right)_{\widetilde{\Gamma}} = 0 \quad \forall w^h \in W^h,\end{aligned} \tag{47}$$

where

$$\left(w^h, 1\right)_{\Omega} = 0, \quad \left(1, P^h u\right)_{\Omega} = (1, u)_{\Omega}. \tag{48}$$

It is assumed under these conditions and subject to a suitably large penalty that the solution of equation (47) is unique (see Elliott and French [3] for details). The problem in equation (47) is in essence the same as the problem

presented in Engel et al. [10], for which error estimates were presented. From these estimates, for $k = 2$,

$$\left\| P^h u - u \right\|_\Omega \leq C h^2 \left\| u \right\|_{3,\Omega}, \tag{49}$$

and for $k > 2$

$$\left\| P^h u - u \right\|_\Omega \leq C h^{k+1} \left\| u \right\|_{k+1,\Omega}. \tag{50}$$

The error in the solution at time t^n is given by

$$c^{h,n} - c\left(t_n\right) = \underbrace{c^{h,n} - P^h c\left(t^n\right)}_{e^{h,n}} + \underbrace{P^h c\left(t^n\right) - c\left(t_n\right)}_{\rho^n}, \tag{51}$$

where we have an estimate for ρ^n, therefore we seek to estimate $e^{h,n}$ in order to bound the error.

From equation (21) and consistency of the formulation,

$$\begin{aligned}
&\left(w, \frac{e^{h,n+1} - e^{h,n}}{\Delta t}\right)_\Omega + \lambda \left(\nabla^2 w, \nabla^2 e^{h,n+\theta}\right)_{\widetilde{\Omega}} + \frac{\alpha\lambda}{h} \left(\left[\!\left[\nabla w\right]\!\right], \left[\!\left[\nabla e^{h,n+\theta}\right]\!\right]\right)_{\widetilde{\Gamma}} \\
&\qquad = -\left(\nabla w^h, \nabla \mu\left(c^{h,n+\theta}\right) - \nabla \mu\left(c\left(t^{n+\theta}\right)\right)\right)_\Omega \\
&\qquad\quad - \left(w, \frac{P^h c\left(t^{n+1}\right) - P^h c\left(t^n\right)}{\Delta t} - c_{,t}\left(t^{n+\theta}\right)\right)_\Omega \\
&\qquad - \lambda \left(\nabla^2 w^h, \nabla^2 \left((1-\theta)\, c\left(t^n\right) + \theta c\left(t^{n+1}\right) - c\left(t^{n+\theta}\right)\right)\right)_{\widetilde{\Omega}} \\
&\qquad + \lambda \left(\left[\!\left[\nabla w^h\right]\!\right], \left\langle \nabla^2 e^{h,n+\theta} \right\rangle\right)_{\widetilde{\Gamma}} + \lambda \left(\left\langle \nabla^2 w^h \right\rangle, \left[\!\left[\nabla e^{h,n+\theta}\right]\!\right]\right)_{\widetilde{\Gamma}} \\
&\qquad + \lambda \left(\left[\!\left[\nabla w^h\right]\!\right], \left\langle \nabla^2 \left((1-\theta)\, c\left(t^n\right) + \theta c\left(t^{n+1}\right) - c\left(t^{n+\theta}\right)\right)\right\rangle\right)_{\widetilde{\Gamma}}.
\end{aligned} \tag{52}$$

We set $w^h = e^{h,n+\theta}$ in this relation, and consider α sufficiently large such that $\exists\ C^\star > 0$ such that $\forall\ w^h \in W^h$

$$\begin{aligned}
&\left(\nabla^2 w^h, \nabla^2 w^h\right)_{\widetilde{\Omega}} - 2\left(\left[\!\left[\nabla w^h\right]\!\right], \left\langle \nabla^2 w^h \right\rangle\right)_{\widetilde{\Gamma}} + \frac{\alpha}{h} \left(\left[\!\left[\nabla w^h\right]\!\right], \left[\!\left[\nabla w^h\right]\!\right]\right)_{\widetilde{\Gamma}} \\
&\qquad\qquad \geq C^\star \left(\left\| \nabla^2 w^h \right\|^2_{\widetilde{\Omega}} + \frac{\alpha}{h} \left\| \left[\!\left[\nabla w^h\right]\!\right] \right\|^2_{\widetilde{\Gamma}}\right)
\end{aligned} \tag{53}$$

(which is effectively stability of the formulation). This leads to

$$\begin{aligned}
&\left(e^{h,n+\theta}, \frac{e^{h,n+1} - e^{h,n}}{\Delta t}\right)_\Omega + C^\star \lambda \left\| \nabla^2 e^{h,n+\theta} \right\|^2_{\widetilde{\Omega}} + C^\star \frac{\alpha\lambda}{h} \left\| \left[\!\left[\nabla e^{h,n+\theta}\right]\!\right] \right\|^2_{\widetilde{\Gamma}} \\
&\qquad \leq \left\| \mu\left(c^{h,n+\theta}\right) - \mu\left(c\left(t^{n+\theta}\right)\right) \right\|_\Omega \left\| \nabla^2 e^{h,n+\theta} \right\|_{\widetilde{\Omega}} \\
&\qquad\quad + \frac{C}{h} \left\| \mu\left(c^{h,n+\theta}\right) - \mu\left(c\left(t^{n+\theta}\right)\right) \right\|_\Omega \left\| \left[\!\left[\nabla e^{h,n+\theta}\right]\!\right] \right\|_{\widetilde{\Gamma}} \\
&\qquad\quad + \left\| \frac{P^h c\left(t^{n+1}\right) - P^h c\left(t^n\right)}{\Delta t} - c_{,t}\left(t^{n+\theta}\right) \right\|_\Omega \left\| e^{h,n+\theta} \right\|_\Omega \\
&\qquad + \lambda \left\| \nabla^2 \left((1-\theta)\, c\left(t^n\right) + \theta c\left(t^{n+1}\right) - c\left(t^{n+\theta}\right)\right) \right\| \left\| \nabla^2 e^{h,n+\theta} \right\|_{\widetilde{\Omega}} \\
&\quad + \lambda \left\| \left[\!\left[\nabla e^{h,n+\theta}\right]\!\right] \right\|_{\widetilde{\Gamma}} \left\| \left\langle \nabla^2 \left((1-\theta)\, c\left(t^n\right) + \theta c\left(t^{n+1}\right) - c\left(t^{n+\theta}\right)\right)\right\rangle \right\|_{\widetilde{\Gamma}}.
\end{aligned} \tag{54}$$

Using equation (30), the above can be rearranged such that

$$\left(e^{h,n+\theta}, \frac{e^{h,n+1}-e^{h,n}}{\Delta t}\right)_{\Omega} + C^{\star}\lambda \left\|\nabla^2 e^{h,n+\theta}\right\|^2_{\widetilde{\Omega}} + C^{\star}\frac{\alpha\lambda}{h}\left\|\left[\!\left[\nabla e^{h,n+\theta}\right]\!\right]\right\|^2_{\widetilde{\Gamma}}$$
$$\leq \left[C\left\|\mu\left(c^{h,n+\theta}\right) - \mu\left(c\left(t^{n+\theta}\right)\right)\right\|_{\Omega} + \left\|\frac{P^h c\left(t^{n+1}\right) - P^h c\left(t^n\right)}{\Delta t} - c_{,t}\left(t^{n+\theta}\right)\right\|_{\Omega}\right.$$
$$\left. + (1+C)\,\lambda\left\|\nabla^2\left((1-\theta)\,c\left(t^n\right) + \theta c\left(t^{n+1}\right) - c\left(t^{n+\theta}\right)\right)\right\|_{\Omega}\right]$$
$$\left(\left\|\nabla^2 e^{h,n+\theta}\right\|_{\widetilde{\Omega}} + \frac{1}{h}\left\|\left[\!\left[\nabla e^{h,n+\theta}\right]\!\right]\right\|_{\widetilde{\Gamma}}\right). \quad (55)$$

In this form, together with the results in equations (49) and (50), the analysis of Elliott and French [3, Theorem 3.1] can be applied directly, yielding the estimate

$$\left\|c^{h,n} - c(t_n)\right\|_{\Omega} \leq C\left(h^p + \Delta t^q\right), \quad (56)$$

where C is dependent on the exact solution. For $k = 2$ gives $p = 2$, and $k > 2$ gives $p = k + 1$. For $\theta \in (1/2, 1]$ leads to $q = 1$ and $\theta = 1/2$ leads to $q = 2$.

5 Numerical example of phase separation

The numerical example presented in this section illustrates a response which is typical for the Cahn-Hilliard equation. Phase separation is modelled on a unit square from an initially uniform state which is randomly perturbed. The parameters for the example are given in non-dimensional form. Consider therefore a length scale L_0, which is representative of the size of the domain Ω, and time scale $T_0 = L_0^4/D\lambda$. Relevant dimensionless quantities, denoted with an asterisk, are given by:

$$t^{\star} = t/T_0, \quad x^{\star} = x/L_0, \quad \mu_c^{\star} = \mu_c L_0^2/\lambda. \quad (57)$$

Using these, the dimensionless counterpart of equation (1) is given by:

$$c_{,t^{\star}} = \nabla^{\star}\cdot\beta\nabla^{\star}\left(\mu_c^{\star} - \nabla^{\star 2}c\right), \quad (58)$$

where β is a dimensionless term reflecting the nature of the mobility. In the case of constant mobility $\beta = 1$, and in case of degenerate mobility $\beta = c\,(1-c)$.

For this test, the following parameters have been adopted: $\omega/kT = 3$, $NkTL_0^2/\lambda = 3000$, $\beta = c\,(1-c)$ (degenerate mobility), $\alpha = 5$ and $\Delta t^{*} = 2\times 10^{-8}$. For the initial conditions, the average concentration is equal to 0.63, with random fluctuations of zero mean and no fluctuation greater than 0.05. Triangular elements with quadratic basis functions ($k = 2$) have been used and the problem has been stepped in time using the Crank-Nicolson method and a full Newton procedure.

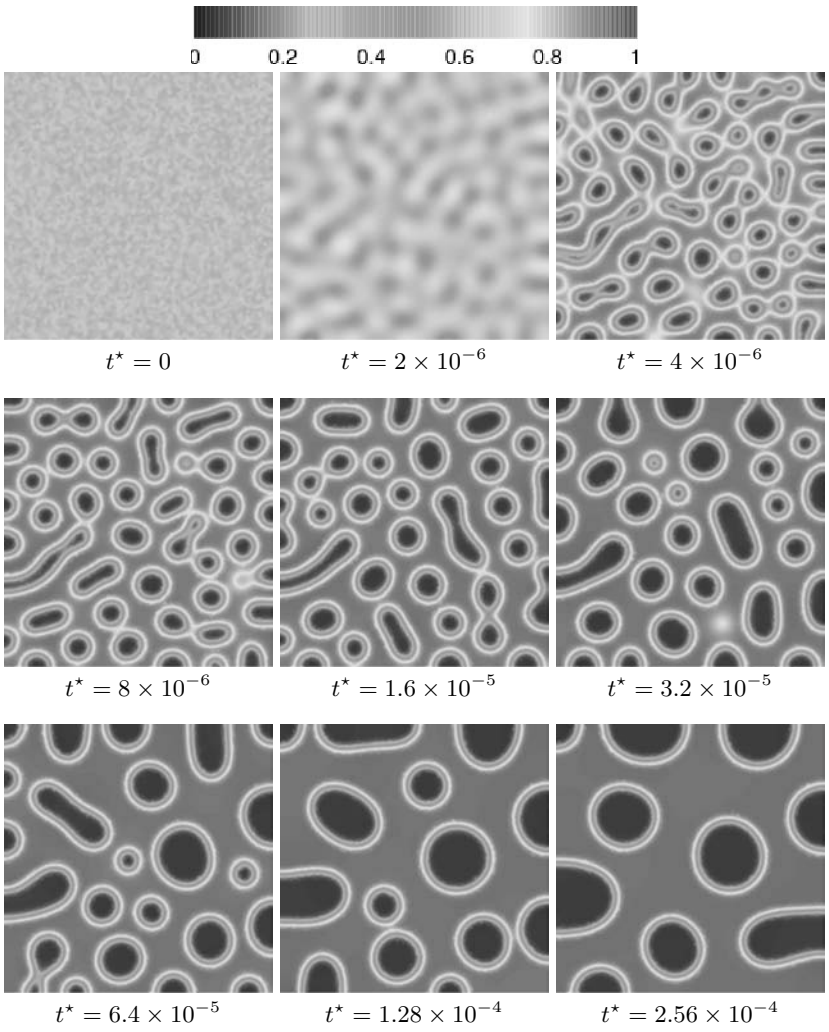


Fig. 1. Evolution of concentration contours from a randomly perturbed initial condition.

Using a random triangulation with $h \approx 1/100$, the evolution of the concentration field is depicted in Figure 1. The concentration evolution can basically be categorised in two phases: the first phase, which is predominantly governed by spinodal decomposition and phase separation, and a second phase which is characterised by grain coarsening. During the first phase, changes in concentration are driven primarily by the minimisation of the local chemical energy Ψ^c. This period is basically terminated as soon as the local concentration is driven to either value of the two binodal points. Approximately from

$t^\star = 8 \times 10^{-6}$ onwards, local changes in concentrations are primarily governed by the surface free energy Ψ^s. In order to minimise its contribution, the generated patterns cluster and grains tend to coarsen. This Ostwald ripening takes place at a much longer time scale.

6 Conclusions

A fully discrete finite element formulation for the Cahn-Hilliard equation has been analysed. The formulation requires consideration of the concentration field only and exploits simple Lagrange finite element basis functions. The necessary continuity of derivatives across element boundaries is enforced in a weak sense and a penalty term acting on jumps in the normal derivative across element boundaries is added to maintain stability.

It is shown for Crank-Nicolson and backward Euler time stepping schemes that critical time steps for both uniqueness and stability exist, but are independent on the element size h. Interestingly, the allowable time step is related to the penalty term. A larger penalty term enhances stability and allows for a larger time step. It is possible that the presented time step restrictions could be tightened and the time step restriction for uniqueness quantified. For the explicit forward Euler scheme, the critical time step scales with h^4, making the scheme impractical. Finally, the fully discrete problem has been posed in such a form that an existing *a priori* error estimate for the Cahn-Hilliard equation in the L^2 norm can be applied directly, proving optimal rates of convergence for the proposed scheme.

Acknowledgements

GNW was supported by the Netherlands Technology Foundation STW, applied science division of NWO and the technology programme of the Ministry of Economic Affairs. KG was supported by a Presidential Early Career Award for Scientists and Engineers through the US Department of Energy.

References

1. Cahn JW, Hilliard JE (1958) Free energy of a nonuniform system-I: Interfacial free energy. The Journal of Chemical Physics 28(2):258–267.
2. Elliott CM, Zheng S (1986) On the Cahn-Hilliard Equation. Archive for Rational Mechanics and Analysis 96(4):339–357.
3. Elliott CM, French DA (1987) Numerical studies of the Cahn-Hilliard equation for phase separation. IMA Journal of Applied Mathematics 38:97–128.

4. Elliott CM, French DA (1989) A non-conforming finite element method for the two-dimensional Cahn-Hilliard equation. SIAM Journal on Numerical Analysis 26(4):884–903.
5. Elliott CM, French DA, Milner FA (1989) A 2nd-order splitting method for the Cahn-Hilliard equation. Numerische Mathematiek 54(5):575–590.
6. Du Q, Nicolaides RA (1991) Numerical analysis of a continuum model of phase transition. SIAM Journal on Numerical Analysis 28(5):1310–1322.
7. Copetti MIM, Elliot CM (1992) Numerical analysis of the Cahn-Hilliard equation with a logarithmic free energy. Numerische Mathematik 63(1):39–65.
8. Blowey JF, Elliott CM (1992) The Cahn-Hilliard gradient theory for phase separation with non-smooth free energy Part II: Numerical analysis. European Journal of Applied Mathematics 3:147–179.
9. Wells GN, Kuhl E, Garikipati K (2006) A discontinuous Galerkin formulation for the Cahn-Hilliard equation. Journal of Computational Physics 218(2):860–877.
10. Engel G, Garikipati K, Hughes TJR, Larson MG, Mazzei L, Taylor RL (2002) Continuous/discontinuous finite element approximations of fourth-order elliptic problems in structural and continuum mechanics with applications to thin beams and plates, and strain gradient elasticity. Computer Methods in Applied Mechanics and Engineering 191(34):3669–3750.

FINITE ELEMENT METHODS WITH EMBEDDED DISCONTINUITIES

Recent Developments in the Formulation of Finite Elements with Embedded Strong Discontinuities

Francisco Armero* and Christian Linder

Department of Civil and Environmental Engineering,
University of California, Berkeley CA 94720-1710, USA

Summary. This paper presents new finite elements with embedded strong discontinuities for the modeling of failure in solids. The new elements consider linear interpolations of the displacement jumps, leading to approximations free of stress locking in general displacement, mixed, assumed and enhanced strain quadrilateral elements. The new interpolations allow also the formulation of globally continuous interpolations along the discontinuity path as developed in this paper too. Several numerical examples are presented to illustrate the performance of the new formulations.

Key words: Strong discontinuities, finite elements, failure of solids.

1 Introduction

The finite element modeling of propagating discontinuities has received a great deal of attention recently, especially for the modeling of the failure of solids in the form of cracks or similar. A main challenge has been the modeling of strong discontinuities (or discontinuities in the displacements) independently of the mesh, thus avoiding any remeshing. Two main approaches have been proposed in this respect: the earlier approach based on the local enhancement of the finite elements [20; 5; 6; 17; 9; 14] and the nodally based enrichment in partition of unity methods [7; 15; 23; 24], among others.

The consideration of enhanced elements has the advantage of allowing the static condensation of the local parameters modeling the strong discontinuities at the element level. Their foundation can be traced to a multi-scale characterization of these solutions by which the standard mechanical boundary-value problem with standard regularity conditions defines the large-scale problem, incorporating the localized dissipative effects of the discontinuities locally through the consideration of vanishing small scales; see [2; 3] for details. This

* Corresponding author (armero@ce.berkeley.edu).

Alain Combescure et al. (eds.), IUTAM Symposium on Discretization Methods for Evolving Discontinuities, 105–122.

framework allowed us the formulation of finite elements with embedded strong discontinuities for infinitesimal [5] and finite deformation continuum problems [6; 2], coupled poroplasticity [10] and beam/plate problems [11; 4]. Alternative treatments include the regularized discontinuity approach considered in [17; 12] and the enhanced finite elements presented in [22; 9; 16], among others.

Many of these works consider triangular finite elements with piece-wise constant interpolations of the discontinuity jumps. Quadrilateral elements with constant jumps have been considered in [9; 16]. Constant jumps, however, proved to be inappropriate for the kinematics of beam and plates [11; 4], as the so-called stress locking appears. This situation can be defined as the overstiff response of the elements due to a spurious transfer of stresses through the discontinuity, leading in particular to the impossibility to reproduce a fully open discontinuity with no associated stresses (i.e. in combination with rigid body modes of the separating parts of the element). It was observed in [11; 4] that the consideration of this precise modes allowed the design of locking-free enhancements, including linear interpolations of the deflections in that context. A similar approach has been recently followed in [14] for the formulation of quadrilateral elements with regularized discontinuities in the continuum, since these elements exhibit a similar spurious transfer of stresses when trying to model non-constant separation modes.

We develop in this paper elements with linear interpolations of the jumps, exploiting again the approach of identifying the strain modes to be captured exactly. This approach contrasts to a direct description of the discontinuous jumps over the element as followed in [1; 8]. We consider general quadrilateral elements, displacement-based as well as mixed, assumed and enhanced strain elements. In contrast with the elements presented in [14], we consider exact non-regularized discontinuities and with linear interpolations of not only the normal jump but also the tangential component. The later requires not only the capture of the rigid body modes of the splitting element, but also a discontinuous relative stretching of the different parts.

Furthermore, the availability of fully linear jumps allows the enforcement of the continuity of these interpolations along the global discontinuity path, a second new outcome of this paper, with a vanishing opening displacement at the discontinuity tip. Remarkably, this continuous interpolation can be accomplished while still maintaining the aforementioned local implementation by which the enhanced parameters modeling the jumps are statically condensed at the element level, and hence implying minimal changes to a standard finite element code. This situation is to be contrasted with available approaches like in [1; 8] based on the addition of nodes along the discontinuity path leading to global enrichments.

An outline of the rest of the paper is as follows. Section 2 presents a summary of the theoretical modeling of strong discontinuities, and its approximation through enhanced finite elements. The new elements with linear jumps are developed in Section 3, with the variation imposing their global

continuity presented in Section 4. The performance of these new formulations is evaluated in Section 5 with several representative numerical simulations. The paper concludes with several final remarks in Section 6.

2 Strong discontinuities and their finite element approximation

After presenting the equations defining the problem of interest in this paper in Section 2.1 (namely, a multi-scale framework of strong discontinuities for the modeling of failure in solids), we present in Section 2.2 the general finite element framework considered in the development of the new elements proposed in this work.

2.1 Problem definition

We are interested in analyzing the failure of a solid $\Omega \subset \mathbb{R}^{n_{dim}}$ whose deformation can be characterized in the infinitesimal range of interest by a displacement field $\boldsymbol{u} : \Omega \rightarrow \mathbb{R}^{n_{dim}}$ in the large scale. By this we consider that this field exhibits the standard regularity conditions (e.g. $H^1(\Omega)$ typical of elastic problems), defining the static equilibrium in the large scale by the standard weak relation

$$\int_{\Omega} \boldsymbol{\sigma} : \nabla^s \boldsymbol{\omega} \, d\Omega = \int_{\Omega} \rho \boldsymbol{b} \cdot \boldsymbol{\omega} \, d\Omega + \int_{\partial_T \Omega} \bar{\boldsymbol{T}} \cdot \boldsymbol{\omega} \, dS \, , \tag{1}$$

for all admissible variations $\boldsymbol{\omega}$ of $\boldsymbol{u}$ (that is, vanishing on the part of the boundary $\partial_u \Omega$ where the displacement field $\boldsymbol{u}$ is imposed), and for the given applied volumetric loading $\rho \boldsymbol{b}$ and tractions $\bar{\boldsymbol{T}}$ on part of the boundary $\partial_T \Omega$.

The stresses $\boldsymbol{\sigma}$ in (1) are given in terms of the strains $\boldsymbol{\varepsilon}(\boldsymbol{u}) = \nabla^s \boldsymbol{u}$ and other effects characteristic of the solid's response. Here we are interested in the modeling of the failures that can be characterized by a cohesive law along surfaces of discontinuity of the displacements, the so-called strong discontinuities. In the multi-scale framework outlined here, these discontinuities are considered locally, as a surface Γ_x with normal $\boldsymbol{n}$ defined locally in a neighborhood $\Omega_x \subset \Omega$ of the points where its appearance has been detected (e.g. acoustic tensor condition on the bulk response). The local displacements are then written in the form

$$\boldsymbol{u}_\mu = \boldsymbol{u} + \tilde{\boldsymbol{u}}(\llbracket \boldsymbol{u}_\mu \rrbracket) \quad \text{in } \Omega_x \, , \tag{2}$$

that is, through a discontinuous displacement field $\tilde{\boldsymbol{u}}$ added to the original (regular) large-scale displacement $\boldsymbol{u}$ and defined in terms of the associated displacement jumps $\llbracket \boldsymbol{u}_\mu \rrbracket : \Gamma_x \rightarrow \mathbb{R}^{n_{dim}}$ across Γ_x. Assuming for example a linear elastic response of the bulk $\Omega_x \backslash \Gamma_x$, the stresses are given by

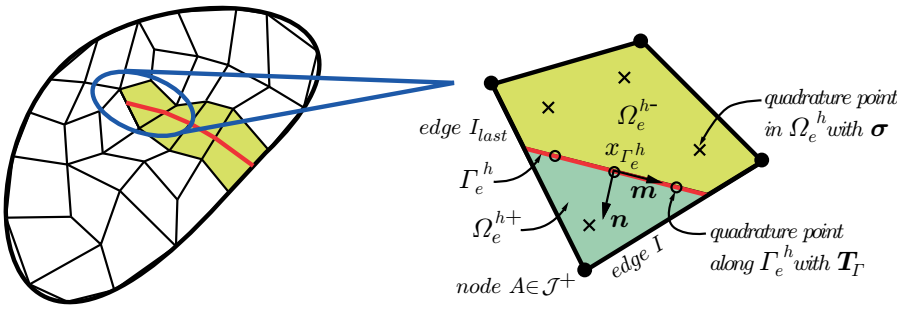


Fig. 1. Typical finite element mesh with a strong discontinuity propagating through it enhancing the elements along its path.

$$\boldsymbol{\sigma} = \mathbb{C}\boldsymbol{\varepsilon}_\mu \quad \text{for} \quad \boldsymbol{\varepsilon}_\mu = \boldsymbol{\varepsilon}(\boldsymbol{u}) + \tilde{\boldsymbol{\varepsilon}}(\llbracket \boldsymbol{u}_\mu \rrbracket) \quad \text{in } \Omega_x \backslash \Gamma_x \,, \tag{3}$$

depending, in addition to the large-scale strains $\boldsymbol{\varepsilon}(\boldsymbol{u})$, on the enhanced strains $\tilde{\boldsymbol{\varepsilon}}(\llbracket \boldsymbol{u}_\mu \rrbracket)$ giving the effects of the jumps $\llbracket \boldsymbol{u}_\mu \rrbracket$ along Γ_x to the rest of Ω_x.

The newly introduced local fields $\llbracket \boldsymbol{u}_\mu \rrbracket$ on Γ_x are determined through the local equilibrium relation across Γ_x

$$\int_{\Gamma_x} \llbracket \boldsymbol{\omega}_\mu \rrbracket \cdot (\boldsymbol{\sigma}\boldsymbol{n} - \boldsymbol{T}_\Gamma) \; dS = 0 \,, \tag{4}$$

for all variations $\llbracket \boldsymbol{\omega}_\mu \rrbracket$ of the jumps $\llbracket \boldsymbol{u}_\mu \rrbracket$ and for the driving traction $\boldsymbol{T}_\Gamma$ defining the aforementioned cohesive law $\boldsymbol{T}_\Gamma(\llbracket \boldsymbol{u}_\mu \rrbracket)$ (i.e. a softening law) that introduces the localized dissipative mechanism to be captured (e.g. a crack). We refer to [2; 3] for details on all these considerations and, in particular, the analysis of the existence of solutions in the so-called "large-scale limit" $h_x := measure(\Omega_x)/measure(\Gamma_x) \to 0$ of vanishing small scales Ω_x, the case of interest here.

2.2 The finite element approximation

The most attractive feature of the formulation outlined above for the treatment of strong discontinuities is its local character. It directly translates to the numerical setting by identifying the neighborhood Ω_x with a finite element Ω_e^h (note also the consistency of these considerations with the notion of the limit $h_x \to 0$ above). We then consider the discontinuities propagating through a general mesh, defined by straight segments Γ_e^h over each active element for the plane problems considered here. See Figure 1 for an illustration. It remains then to define the discrete counterparts of the local fields identified in the previous section, for *a single finite element* once its local geometry with Γ_e^h is known.

We start by considering a local interpolation of the displacement jumps along the segment Γ_e^h of the form

$$[\![\boldsymbol{u}_{\mu}^{h}]\!](s) = \boldsymbol{J}(s)\ \boldsymbol{\xi}\ , \tag{5}$$

for a set of local element parameters $\boldsymbol{\xi}$ and associated interpolation functions $\boldsymbol{J}(s)$ defined in terms of a local coordinate s in Γ_e^h. The main aspect of the formulation developed in the previous section for the continuum problem is the effect of the jumps (5) on the strains in the bulk $\Omega_e^h \backslash \Gamma_e^h$ as considered in (3). In the (infinitesimal) discrete setting of interest here, we have

$$\boldsymbol{\varepsilon}_{\mu}^{h} = \bar{\boldsymbol{B}}\boldsymbol{d} + \boldsymbol{G}_{(c)}\boldsymbol{\xi} \qquad \text{in} \quad \Omega_e^h \backslash \Gamma_e^h\ , \tag{6}$$

for an enhanced strain operator $\boldsymbol{G}_{(c)}$ to be defined over the element Ω_e^h, identifying the contributions of the discontinuity to the discrete kinematics of the underlying element. This kinematics is defined in (6), as usual, through a set of nodal displacements $\boldsymbol{d}$, interpolating the large-scale displacement $\boldsymbol{u}^h = \boldsymbol{N}\boldsymbol{d}$ for a set of shape functions $\boldsymbol{N}$, but with a general discrete strain operator $\bar{\boldsymbol{B}}$ for the large-scale strains $\boldsymbol{\varepsilon}^h = \bar{\boldsymbol{B}}\boldsymbol{d}$ to accommodate general finite element methods like assumed strain, mixed or enhanced strain formulations.

In this discrete setting, the governing equations read

$$\boldsymbol{R} = \boldsymbol{f}_{ext} - \mathop{\mathbf{A}}_{e=1}^{n_{elem}} \left(\int_{\Omega_e^h} \bar{\boldsymbol{B}}^T \boldsymbol{\sigma}\ d\Omega \right) = 0\ , \tag{7}$$

$$\boldsymbol{r}_{enh}^{e} = -\int_{\Omega_e^h} \boldsymbol{G}_{(e)}^T \boldsymbol{\sigma}\ d\Omega - \int_{\Gamma_e^h} \boldsymbol{J}^T \boldsymbol{T}_{\Gamma}\ d\Gamma = 0\ , \tag{8}$$

where the second equation applies only and separately to the elements Ω_e^h where a discontinuity segment Γ_e^h has been activated. Equation (7) is the discrete counterpart of the global equilibrium equation (1), involving the standard assembly operator $\mathbf{A}$ over the n_{elem} elements for the problem at hand. Similarly, the actual numerical implementation involves the standard Voigt's vector notation, but no explicit distinction in the symbols is made here.

Equation (8) is the discretization of the local equilibrium equation (4). However, in contrast with this last equation, the stress term has been defined with an integral over the element Ω_e^h through the "equilibrium" operator $\boldsymbol{G}_{(e)}$, since the stresses $\boldsymbol{\sigma}$ are naturally defined there. In fact, they are evaluated at the quadrature points employed in the evaluation of all the integrals over the element Ω_e^h, as usual. Similarly, the integral over the discontinuity segment Γ_e^h in equation (8) involves quantities defined naturally there (e.g. the driving traction $\boldsymbol{T}_{\Gamma}$ in terms of the local jump $[\![\boldsymbol{u}_{\mu}^{h}]\!]$ and the interpolation functions $\boldsymbol{J}$), being then easily evaluated through a local quadrature rule along Γ_e^h. The implementation developed here considers standard Gauss quadratures for both. See Figure 1 for an illustration of these considerations.

Comparing equations (8) and (4), we see that the equilibrium operator $\boldsymbol{G}_{(e)}$ can be understood as a projection of the stresses defined at the quadrature points in Ω_e^h to the tractions on Γ_e^h for the different components of the assumed interpolations $\boldsymbol{J}$. In this way, we request

$$\int_{\Omega_e^h} \boldsymbol{G}_{(e)}^T \boldsymbol{\sigma} \, d\Omega = - \int_{\Gamma_e^h} \boldsymbol{J}^T \boldsymbol{\sigma} \boldsymbol{n} \, d\Gamma_e^h + l_{\Gamma_e^h} \, o(h_e^{p+1}) \, , \tag{9}$$

for $h_e = A_{\Omega_e^h}/l_{\Gamma_e^h}$ as in h_x above and for $p \geq 0$ (numerical consistency). As developed in [4] in the context of plate problems, the different components of this operator can be constructed with polynomials by noting that the stresses $\boldsymbol{\sigma}$ in the underlying element can also be considered as polynomials up to higher order terms in h_e. As the local interpolation functions $\boldsymbol{J}$ involve also polynomials in the local coordinate s, we can write for the different components s^k ($k = 0, 1$ in the linear elements considered below)

$$\boldsymbol{G}_{(e)_k} \boldsymbol{\eta} = -\frac{1}{h_e} g_{e k}(x, y) \;\; (\boldsymbol{n} \otimes \boldsymbol{\eta})^s \qquad \forall \boldsymbol{\eta} \in \mathbb{R}^{n_{dim}} \qquad \text{for} \quad k = 0, 1 \, , \tag{10}$$

for a set of linear functions in a local Cartesian system $\{x, y\}$ associated with the element of the form

$$g_{e k} = a_k + b_k \; x + c_k \; y \, , \tag{11}$$

for the constants

$$\begin{bmatrix} a_k \\ b_k \\ c_k \end{bmatrix} = \boldsymbol{H}^{-1} \left(\frac{1}{l_{\Gamma_e^h}} \int_{\Gamma_e^h} \begin{bmatrix} s^k \\ s^k \; x \\ s^k y \end{bmatrix} \, d\Gamma \right) \, , \tag{12}$$

all with $k = 0, 1$, and the matrix

$$\boldsymbol{H} := \frac{1}{A_{\Omega_e^h}} \int_{\Omega_e^h} \begin{bmatrix} 1 & x & y \\ x & x^2 & xy \\ y & xy & y^2 \end{bmatrix} d\Omega \, , \tag{13}$$

with $A_{\Omega_e^h} = \int_{\Omega_e^h} d\Omega$ and $l_{\Gamma_e^h} = \int_{\Gamma_e^h} d\Gamma$. The integrals in (12) and (13) are evaluated with the same quadrature rules as described before. They only need to be evaluated once, after the particular geometry of the element crossed by the discontinuity segment Γ_e^h is known.

To complete the finite element formulation, it remains to define the enhanced strain operator $\boldsymbol{G}_{(c)}$ for the particular assumed interpolations $\boldsymbol{J}$ of the displacement jumps. We develop elements enhanced with linear jumps next.

Remark 1. A very important feature of equations (7)-(8) is again that the equations (8) are satisfied at the element level. This fact allows the local elimination of the enhanced parameters $\boldsymbol{\xi}$, thus resulting in the global system of equations

$$\boldsymbol{K}_* \boldsymbol{\Delta d} = \boldsymbol{R}_* \, , \tag{14}$$

for the increments $\boldsymbol{\Delta d}$ of the nodal displacements after the linearization used in the iterative solution of equations (7)-(8). The statically condensed matrix $\boldsymbol{K}_*$ and residual $\boldsymbol{R}_*$ are defined by the relations

$$\boldsymbol{K}_* = \mathop{\mathbf{A}}_{e=1}^{n_{elem}} \left[\boldsymbol{K}^e_{dd} - \boldsymbol{K}^e_{d\xi} \boldsymbol{K}^{e^{-1}}_{\xi\xi} \boldsymbol{K}^e_{\xi d} \right] \ , \ \boldsymbol{R}_* = \mathop{\mathbf{A}}_{e=1}^{n_{elem}} \left[\boldsymbol{R}^e - \boldsymbol{K}^e_{d\xi} \boldsymbol{K}^{e^{-1}}_{\xi\xi} \boldsymbol{r}^e_{enh} \right] \ , \quad (15)$$

for the element matrices

$$\boldsymbol{K}^e_{dd} = \int_{\Omega^h_e} \bar{\boldsymbol{B}}^T \mathbb{C}\, \bar{\boldsymbol{B}} d\Omega \ , \quad \boldsymbol{K}^e_{d\xi} = \int_{\Omega^h_e} \bar{\boldsymbol{B}}^T \mathbb{C}\, \boldsymbol{G}_{(c)} d\Omega \ , \quad (16)$$

$$\boldsymbol{K}^e_{\xi d} = \int_{\Omega^h_e} \boldsymbol{G}^T_{(e)} \mathbb{C}\, \bar{\boldsymbol{B}} d\Omega \ , \quad \boldsymbol{K}^e_{\xi\xi} = \int_{\Omega^h_e} \boldsymbol{G}^T_{(e)} \mathbb{C}\, \boldsymbol{G}_{(c)} d\Omega + \int_{\Gamma^h_e} \boldsymbol{J}^T \tilde{\mathbb{C}}_\Gamma \boldsymbol{J} d\Gamma \ , \quad (17)$$

for the bulk and localized tangents $\mathbb{C}$ and $\tilde{\mathbb{C}}_\Gamma$, respectively (i.e. $\Delta\boldsymbol{\sigma} = \mathbb{C}\Delta\boldsymbol{\varepsilon}^h_\mu$ and $\Delta\boldsymbol{T}_\Gamma = \tilde{\mathbb{C}}_\Gamma \Delta[\![\boldsymbol{u}^h_\mu]\!]$). For future developments, we observe the need to invert $\boldsymbol{K}^e_{\xi\xi}$, a small 4×4 square matrix for the elements developed next with four enhanced parameters $\boldsymbol{\xi}$ per element. We also note the non-symmetric character of the formulation for a general choice $\boldsymbol{G}_{(e)} \neq \boldsymbol{G}_{(c)}$.

3 Finite elements with linear jumps

As discussed in Section 1, most of the finite elements with embedded strong discontinuities developed to date involve constant interpolations of the displacement jumps. This choice, however, may lead to stress locking when combined with quadrilateral and higher-order finite elements, in the sense that spurious stresses are transferred through the discontinuity, even if it is fully softened and no stiffness should arise from it; see [4; 14]. Here we develop elements with linear interpolations of the discontinuity jumps that improve on this problem, especially for general bilinear-based quadrilateral elements.

As developed in [11; 4] for beams and plates respectively, the design of locking-free interpolations can be easily and efficiently accomplished by identifying first the strain modes, say $\widehat{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{\xi})$, to be reproduced exactly by an associated set of nodal displacements $\widehat{\boldsymbol{d}}(\boldsymbol{\xi})$, both in terms of the parameters $\boldsymbol{\xi}$, given the discrete kinematics of the underlying element. That is, in the infinitesimal range of interest here, we impose

$$\widehat{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{\xi}) = \bar{\boldsymbol{B}} \widehat{\boldsymbol{d}}(\boldsymbol{\xi}) + \boldsymbol{G}_{(c)} \boldsymbol{\xi} \qquad \forall \boldsymbol{\xi} \ , \quad (18)$$

identifying the enhanced operator $\boldsymbol{G}_{(c)}$ in the process.

Figure 2 sketches the four modes to be reproduced exactly for the linear interpolations of interest, in the plane setting considered here. In particular, we can find constant and linear separation modes for both the normal and tangential jumps, denoted by $\boldsymbol{\xi} = [\xi_{0n}\ \xi_{0m}\ \xi_{1n}\ \xi_{1m}\ ]^T$, in the local reference system $\{\boldsymbol{n}, \boldsymbol{m}\}$. All these separation modes, with the exception of the fourth one corresponding to the linear tangential separation mode, must be accompanied with no strains in the bulk of the element in order to be able to avoid the stress locking described above. The fourth mode is accompanied with an axial stretching in the direction $\boldsymbol{m}$ on one side of the discontinuity.

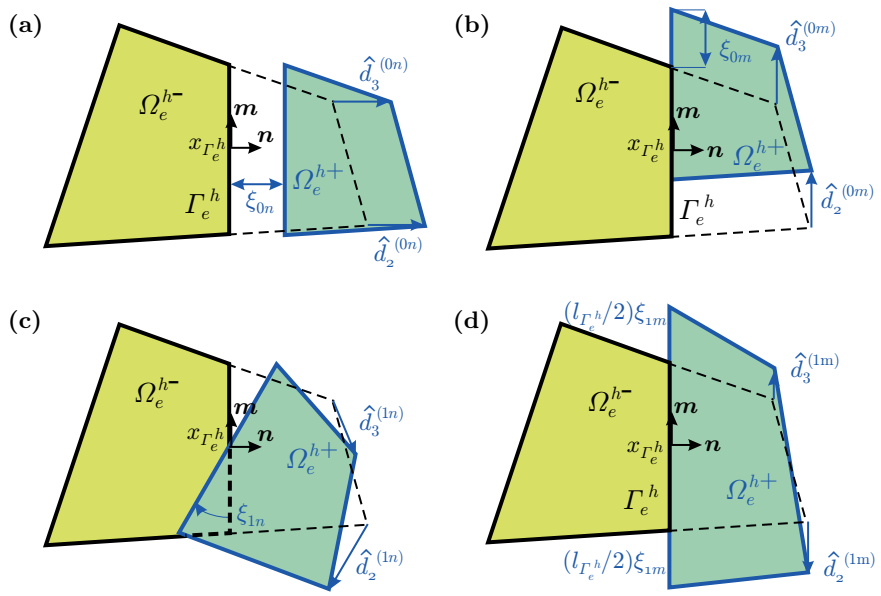


Fig. 2. Strain modes used in the design of the enhanced operator $\boldsymbol{G}_{(c)}$: **(a)** constant normal, **(b)** constant tangential, **(c)** linear normal, and **(d)** linear tangential, separation modes. In all cases, $\mathcal{J}^+ = \{2, 3\}$ for the nodes on the Ω_e^{h+} side.

Proceeding systematically with equation(18) for each mode, we have for the constant modes the nodal displacements

$$\widehat{\boldsymbol{d}}_A^{\langle 0n\rangle}(\boldsymbol{\xi}_0) = \begin{cases} \xi_{0n}\,\boldsymbol{n} & \textit{for } A \in \mathcal{J}^+ \\ 0 & \textit{otherwise} \end{cases} \quad \text{and} \quad \widehat{\boldsymbol{d}}_A^{\langle 0m\rangle}(\boldsymbol{\xi}_0) = \begin{cases} \xi_{0m}\,\boldsymbol{m} & \textit{for } A \in \mathcal{J}^+ \\ 0 & \textit{otherwise} \end{cases} \tag{19}$$

respectively, where $\mathcal{J}^+$ denotes the set of nodes of the part of the element Ω_e^{h+} separated by the discontinuity Γ_e^h; see Figure 1. The choice of one side or the other can be easily proven to be irrelevant in the final formulation. Using equation (18) for the nodal displacements (19) and $\widehat{\boldsymbol{\varepsilon}}_\mu^{\langle 0n\rangle} = \widehat{\boldsymbol{\varepsilon}}_\mu^{\langle 0m\rangle} = 0$, we obtain

$$\boldsymbol{G}_{(c)}^{\langle 0n\rangle} = -\sum_{A\in\mathcal{J}^+} \bar{\boldsymbol{B}}^A \boldsymbol{n} \qquad \text{and} \qquad \boldsymbol{G}_{(c)}^{\langle 0m\rangle} = -\sum_{A\in\mathcal{J}^+} \bar{\boldsymbol{B}}^A \boldsymbol{m} \tag{20}$$

for the enhanced strain operators associated to the constant modes.

The linear normal separation mode follows similarly for $\widehat{\boldsymbol{\varepsilon}}_\mu^{\langle 1n\rangle} = 0$ and the nodal displacements

$$\widehat{\boldsymbol{d}}_A^{\langle 1n\rangle}(\boldsymbol{\xi}_{1n}) = \begin{cases} \xi_{1n} \ (\boldsymbol{n}\otimes\boldsymbol{m})^a \, \bar{\boldsymbol{x}}_A & for\ A \in \mathcal{J}^+ \\ 0 & otherwise \end{cases} \tag{21}$$

corresponding to an infinitesimal rotation of the side $\Omega_e^{h\,+}$ for the scalar enhanced parameter ξ_{1n}, with

$$(\boldsymbol{n}\otimes\boldsymbol{m})^a := \boldsymbol{n}\otimes\boldsymbol{m} - \boldsymbol{m}\otimes\boldsymbol{n}\ , \quad \text{and} \quad \bar{\boldsymbol{x}}_A = \boldsymbol{x}_A - \boldsymbol{x}_{\Gamma_e^h}\ , \tag{22}$$

for the coordinates $\boldsymbol{x}_A$ of nodes $A \in \mathcal{J}^+$ and the center $\boldsymbol{x}_{\Gamma_e^h}$ of the discontinuity segment Γ_e^h, as depicted in Figure 1. Imposing again (18), we obtain

$$\boldsymbol{G}_{(c)}^{\langle 1n\rangle} = -\sum_{A\in\mathcal{J}^+} \bar{\boldsymbol{B}}^A \, (\boldsymbol{n}\otimes\boldsymbol{m})^a \, \bar{\boldsymbol{x}}_A \tag{23}$$

for the enhanced strain operator corresponding to this mode.

Finally, the linear tangential mode is characterized, as discussed above, by the bulk strain

$$\widehat{\boldsymbol{\varepsilon}}_\mu^{\langle 1m\rangle} = \xi_{1m}\,(\boldsymbol{m}\otimes\boldsymbol{m}) \ H_{\Gamma_e^h}\ , \tag{24}$$

and the nodal displacements

$$\widehat{\boldsymbol{d}}_A^{\langle 1m\rangle}(\boldsymbol{\xi}_{1m}) = \begin{cases} \xi_{1n} \ (\boldsymbol{m}\otimes\boldsymbol{m})\,\bar{\boldsymbol{x}}_A & for\ A \in \mathcal{J}^+ \\ 0 & otherwise \end{cases} \tag{25}$$

with $\bar{\boldsymbol{x}}_A$ defined in (23) and the Heaviside function

$$H_{\Gamma_e^h}(\boldsymbol{x}) = \begin{cases} 1 & for\ \boldsymbol{x} \in \Omega_e^{h\,+} \\ 0 & otherwise \end{cases} \tag{26}$$

associated to the discontinuity segment Γ_e^h in element Ω_e^h. Imposing the strain relation (18) for all ξ_{1m}, we obtain

$$\boldsymbol{G}_{(c)}^{\langle 1m\rangle} = (\boldsymbol{m}\otimes\boldsymbol{m}) \ H_{\Gamma_e^h} - \sum_{A\in\mathcal{J}^+} \bar{\boldsymbol{B}}^A \, (\boldsymbol{m}\otimes\boldsymbol{m})\,\bar{\boldsymbol{x}}_A \tag{27}$$

for the enhanced strain operator associated to the linear tangential jumps.

The final enhanced strain operator is then given by

$$\boldsymbol{G}_{(c)} = \left[\boldsymbol{G}_{(c)}^{\langle 0n\rangle} \ \ \boldsymbol{G}_{(c)}^{\langle 0m\rangle} \ \ \boldsymbol{G}_{(c)}^{\langle 1n\rangle} \ \ \boldsymbol{G}_{(c)}^{\langle 1m\rangle}\right], \tag{28}$$

or for as many components of the discontinuity jumps activated for the problem at hand. For example, a particular model of ductile materials may require only tangential slipping but no normal opening.

Remark 2. We observe that for a single node separating on one side (i.e. $card(\mathcal{J}^+) = 1$, as depicted in Figure 1) the linear normal mode (23) can

be written as a linear combination of the enhanced strain operators (20) for the constant modes. This is also the case for the linear tangential mode (27) up to the discontinuous term depending on the Heaviside function $H_{\Gamma_e^h}$. A direct consequence of this situation is that the first term in the expression (17) for the matrix $\boldsymbol{K}^e_{\xi\xi}$ becomes singular, leading necessarily to the singularity of this matrix for the common case of a fully softened discontinuity (i.e. $\tilde{\mathbb{C}}_\Gamma = 0$) and prone to become singular in general. We note the need to invert this matrix during the static condensation process (15). Since this situation is easily detected during the propagation of the discontinuity (i.e. one node separates), this instability can be avoided by activating only the constant modes for that particular element since no stress locking occurs in this particular case; see Section 5.1. However, this option does not lead to an improved performance of the element, nor allows the continuous jumps interpolations developed in Section 4 below. For this reason, we consider for this particular case of a single node separating the alternative (stabilized) residual

$$\boldsymbol{r}^{\langle 1\rangle}_{enh} = -\int_{\Omega_e^h} \boldsymbol{G}^{\langle 1\rangle\, T}_{(e)}\, \boldsymbol{\sigma}\; d\Omega - \int_{\Gamma_e^h} \boldsymbol{J}^{\langle 1\rangle\, T} \boldsymbol{T}_\Gamma\; d\Gamma - \kappa\; \boldsymbol{\xi}_1 = 0\;, \tag{29}$$

for the linear separation modes $\boldsymbol{\xi}_1 = [\xi_{1n}\ \xi_{1m}]^T$ and a stabilization parameter $\kappa > 0$. Equation (29) can be understood as a relaxation of a Lagrange multiplier formulation imposing constant separation jumps, recovering this case in the limit $\kappa \to \infty$. However, here we consider the choice of a finite value for this parameter being controlled only by the stabilization of the corresponding tangent matrix (i.e. $\boldsymbol{K}^e_{\xi\xi} + \kappa \mathbf{1}_2$), since any value $0 < \kappa < \infty$ provides a more accurate solution than the original element with constant jumps, as illustrated in Section 5.1. We note the consistency of this approach since the constant separation modes are not affected. We refer to [13] for complete details of these considerations and to [18] for a discussion on the stability of finite elements with embedded discontinuities.

4 Continuous interpolations of the discontinuity jumps

A big advantage of the new finite elements with linear jumps presented in the previous section over the more standard elements with constant jumps is, besides the added accuracy and locking-free response, the availability now to consider continuous interpolations of the discontinuous jumps along the whole discontinuity Γ^h. As a first approach, this can be easily accomplished by a direct enforcement of the continuity of the jumps at the element edges while still using the finite elements developed in the previous section.

To this purpose, we can write the constraints to be enforced as

$$\boldsymbol{g}^{(I)} := [\![\boldsymbol{u}^h_\mu]\!]^{(I)}_{e+1} - [\![\boldsymbol{u}^h_\mu]\!]^{(I)}_{e} = 0\;, \tag{30}$$

for the edge (I) between consecutive elements Ω^h_e and Ω^h_{e+1} along Γ^h, and by

$$\boldsymbol{g}^{(I_{last})} := -[\![\boldsymbol{u}_{\mu}^{h}]\!]_{e}^{(I_{last})} = 0 \ , \tag{31}$$

for the last edge (I_{last}) enforcing the zero opening at the discontinuity tip. See Figure 1 for an illustration. For the linear elements considered in the previous section, we can write in terms of the jump parameters $\boldsymbol{\xi}^{e}$ for element Ω_{e}^{h}

$$[\![\boldsymbol{u}_{\mu}^{h}]\!]_{e}^{(I)} = \left(\xi_{0n}^{e} + s_{e}^{(I)}\xi_{1n}^{e}\right)\boldsymbol{n}_{e} + \left(\xi_{0m}^{e} + s_{e}^{(I)}\xi_{1m}^{e}\right)\boldsymbol{m}_{e} \ , \tag{32}$$

in the (common) global Cartesian system and the local coordinate $s_{e}^{(I)}$ of edge (I) with the local vectors $\{\boldsymbol{n}^{e}, \boldsymbol{m}^{e}\}$ for the discontinuity segment Γ_{e}^{h} of element Ω_{e}^{h}. The final system of equations reads

$$\boldsymbol{R} := \boldsymbol{f}_{ext} - \mathop{\mathsf{A}}_{e=1}^{n_{elem}} \left(\int_{\Omega_{e}^{h}} \bar{\boldsymbol{B}}^{T}\boldsymbol{\sigma} \, d\Omega \right) = 0 \ , \tag{33}$$

$$\boldsymbol{r}_{enh}^{e} := -\int_{\Omega_{e}^{h}} \boldsymbol{G}_{(e)}^{T}\boldsymbol{\sigma} \, d\Omega - \int_{\Gamma_{e}^{h}} \boldsymbol{J}^{T}\boldsymbol{T}_{\Gamma} \, d\Gamma - \boldsymbol{Q}_{e}^{T}\boldsymbol{\lambda} = 0 \ , \tag{34}$$

$$\boldsymbol{Q}\boldsymbol{\xi} = 0 \ , \tag{35}$$

for the Lagrange multipliers $\boldsymbol{\lambda} = [\lambda^{(I)}]$ (with $I = 1, I_{last}$) enforcing the constraints (35) on all the enhanced parameters $\boldsymbol{\xi}$. Equation (35) corresponds to the original constraints (30)-(31) written as the operators $\boldsymbol{Q}_{e}$ and assembled in $\boldsymbol{Q}$ over the different elements with an active discontinuity segment.

Equation (34) is still written at the element level Ω_{e}^{h} but depends now on the Lagrange multipliers $\lambda^{(I)}$ for the element edges crossed by the local discontinuity segment Γ_{e}^{h} and, hence, common to the elements across those edges. Similarly, equation (35) requires the assembly of the enhanced parameters $\boldsymbol{\xi}$ on all these elements thus destroying, in principle, the local structure exploited in Remark 1 in the static condensation of these parameters.

However, this situation can be avoided by the proper relaxation of the constraints (35) and their enforcement through an iterative procedure, that is, an Uzawa-type scheme. In this way, we approach the solution of the equations (33)-(34) in a given time step of the global problem by the solution first of

$$\boldsymbol{R}(\boldsymbol{d}^{l+1}, \boldsymbol{\xi}^{l+1}) = 0 \ , \tag{36}$$

$$\boldsymbol{r}_{enh}^{e}(\boldsymbol{d}^{l+1}, \boldsymbol{\xi}^{l+1}, \boldsymbol{\lambda}^{l}) = 0 \ , \tag{37}$$

for a fixed value of the Lagrange multipliers $\boldsymbol{\lambda}^{l}$ (with, say, $\boldsymbol{\lambda}^{0}$ = value for last step or zero), followed by the local update

$$\boldsymbol{\lambda}^{l+1} = \boldsymbol{\lambda}^{l} + \varrho \, \boldsymbol{Q}\boldsymbol{\xi}^{l+1} \ , \tag{38}$$

for a relaxation parameter $\varrho > 0$.

The system of equations (36)-(37) has the same form as the original system of equations (7)-(8) except for the added term in (34) depending on the fixed

values $\boldsymbol{\lambda}^l$ readily available at the element level. Hence, its solution can still be accomplished with the very same static condensation procedure outlined in Remark 1, involving only local (element) modifications before the assembly of the global system of equations (14). Furthermore, the explicit character of the update (38) makes it very simple to evaluate as the iteration in l proceeds.

Therefore, we have accomplished the enforcement of the continuity constraints with very minor modifications of the code used in the original element with discontinuous interpolations, already a simple modification of a general finite element code given the local nature of the enhancements. Clearly, this very efficient structure of the numerical implementation has been obtained at the price of the stability of the iterative procedure (36)-(38) for large values of the relaxation parameter ϱ. Nevertheless, these simple modifications allow a first and preliminary evaluation of the performance of this type of continuous jump interpolations along the discontinuity path, as we investigate numerically in Section 5.4 below.

5 Representative numerical simulations

We present in this section several numerical simulations to illustrate the performance of the newly proposed elements. Of the main interest is the confirmation of their locking-free character.

5.1 Element bending test

We first consider the bending test [14] consisting of a $200 \times 200 \times 1$ mm^3 block first pulled uniformly across the height (so a discontinuity opens in Mode I through the whole block upon reaching the tensile strength), followed by a double rate of pulling at the bottom with respect to the top (so the discontinuity opens with a linear distribution of the normal jump); see the insets in Figure 3. The bulk is linear elastic with Young modulus 30 GPa and Poisson ratio 0.2, and with a tensile strength of 3 MPa. A linear softening law with modulus -45 MPa/mm is assumed between the normal opening and traction across the discontinuity until the later softens completely to zero. We refer to [13] for the analytical exact solution for this case, including the nonlinear part as the fully softened zone propagates along the discontinuity.

Figure 3 depicts the solution obtained with a single square element for the whole block. We consider the enhanced Q1/E4 element of [21] through the proper $\bar{\boldsymbol{B}}$ operator as the underlying element (an element able to reproduce exactly a pure bending state of stress in the considered regular configuration). As shown in the load-displacement curve of Figure 3, the new finite elements with linear jumps are able to capture the exact solution (it is actually exact during the linear softening part). This situation is to be contrasted with the solution obtained with the original elements with constant jumps. The solution is not only less accurate during the softening solution in this case, but also

BENDING TEST

Single square element **Two distorted elements**

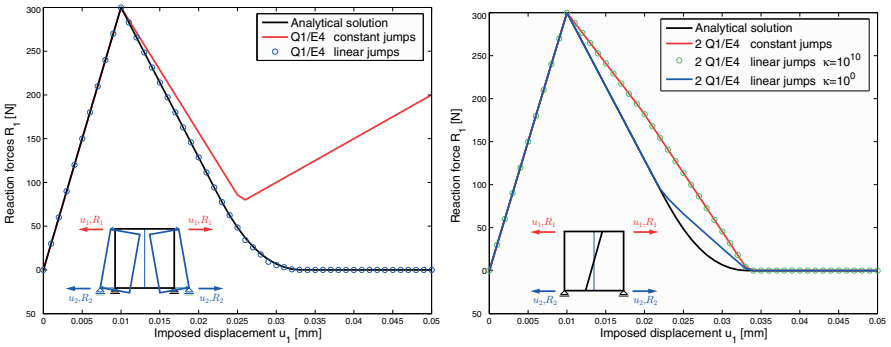


Fig. 3. Element bending test. (Left) Single element under bending with a linearly opening discontinuity showing the stress locking of the elements with constant jumps while the new elements resolve the analytical solution. (Right) Same solution resolved with two distorted elements with a single node separating: the new elements with linear jumps show an improved performance for any finite stabilization parameter κ, recovering in the limit $\kappa \to \infty$ the response of the elements with constant jumps not locking in this case. The underlying continuum element is the enhanced Q1/E4 element in both cases.

PARTIAL TENSION TEST

Single Q1 element **Single Q2 element**

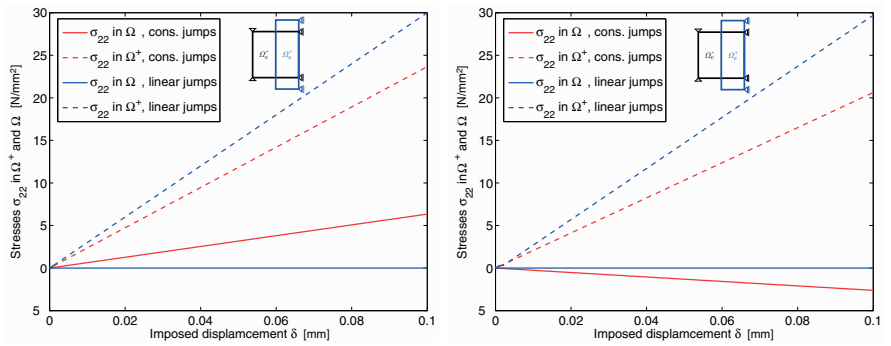


Fig. 4. Partial tension element test. Solutions obtained by the new enhanced elements with linear jumps and the original elements with constant jumps for the 4-node bilinear (Q1) and 9-node quadratic (Q2) quads. The stress on both sides of the fully open discontinuity is shown for the applied displacements, illustrating the spurious transfer of stress through the discontinuity for the elements with constant jumps, while no stresses are transferred for the new elements with linear jumps.

the discontinuity cannot be fully softened, showing a stiff hardening branch due to the spurious transfer of stress across the discontinuity.

Figure 3 includes also the solution obtained with two Q1/E4 distorted elements in such a way that a single node separates for each element. We observe that the elements with constant jumps do not lock in this particular case, but still show a poor solution in the softening branch. The elements with linear jumps show the singularity described in Remark 2 for this particular configuration. We observe that the stabilization procedure (29) is able to obtain a good solution for a small value of the stabilization parameter κ to avoid the singularity, recovering effectively the response of the elements with constant jumps for large values. As noted in that remark, the choice for this parameter is simply dictated by the avoidance of the singularity, obtaining a better performance than the original elements for any finite value.

5.2 Element partial tension test

To evaluate the performance of the elements in the presence of linear tangential jumps, we consider the test sketched in the insets of Figure 4. A straight fully softened discontinuity with no jumps, is set initially through the same square block considered in the previous section. The nodes on one side of the discontinuity are then pulled in the direction of the discontinuity, as shown in Figure 4, creating an axial stretch on that side while no strain (or stress) should be transferred to the other side of the fully softened discontinuity.

Figure 4 shows the solutions obtained for a bilinear 4-node Q1 and a quadratic 9-node Q2 element with constant and linear jumps for the discontinuity. The elements with linear jumps are able to avoid any transfer of stresses through the discontinuity thanks to the new linear tangential mode, in contrast with the original elements with constant jumps.

5.3 Three-point bending test

We consider this classical benchmark problem for a $2000 \times 200 \times 50\ mm^3$ simply supported beam with a $20 \times 100 \times 50\ mm^3$ notch at its bottom center, while loaded by an imposed displacement at the top center; see Figure 5. The Young modulus is 30 GPa, the Poisson ratio 0.2, and the tensile strength 3.33 MPa. A crack propagates from the notch in Mode I. We consider the power law in [19] with a fracture energy 0.124 N/mm for the cohesive law in the normal opening, and a retention factor of 10^{-3} for the tangential direction.

Figure 5 depicts the solution obtained by the new finite elements with linear jumps based on the bilinear Q1 quad, showing a zoom of the notch with the elements having active enhanced parameters. The computed load-displacement curves are presented in Figure 6, comparing several cases. The left plot in this figure compares the performance of this enhanced Q1 element with the similar Q1 element but with constant jumps only, and for two different meshes. A good agreement with the experimental envelope can be observed

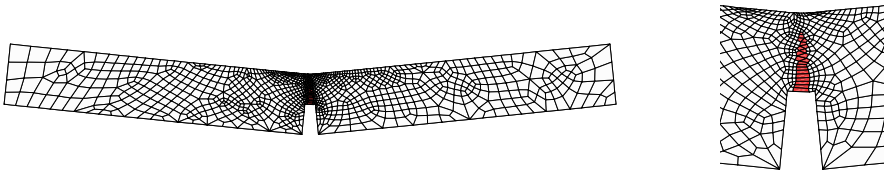


Fig. 5. Three-point bending test. Solution obtained by the new elements with linear jumps and a mesh with 1,105 quadrilateral Q1 enhanced elements, with a zoom of the elements with an active discontinuity shown in red (deformations scaled by 100).

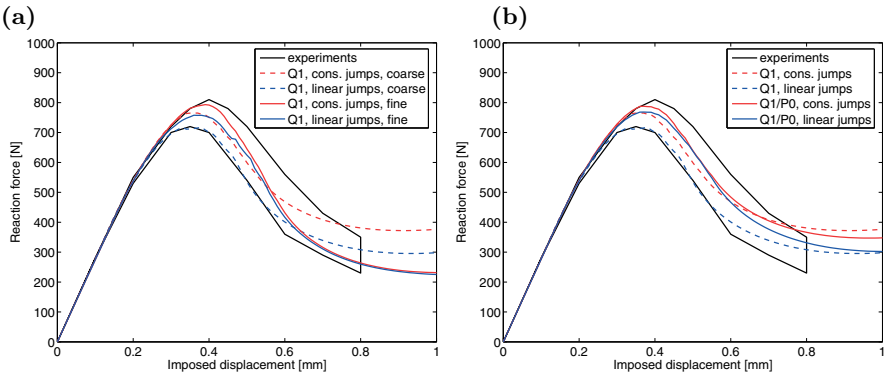


Fig. 6. Three-point bending test. Computed load-displacement curves for elements enhanced with linear and constant jumps. **(a)** Comparison for two different meshes, the mesh above with 1,105 quadrilateral elements and a finer mesh with 2,546 elements. **(b)** Comparison of the solutions obtained with enhanced Q1 and Q1/P0 elements for the coarser mesh above.

in all cases, with the elements with linear jumps showing systematically an improved softer performance, especially in the later stages of deformation.

Figure 6 includes also the load-displacement curves obtained with mixed B-bar Q1/P0 elements enhanced with constant and linear jumps. This illustrates the flexibility of the proposed formulation when considering any particular element and mesh for the resolution of the large-scale problem. Again, the elements with linear jumps show a systematic softer performance, as desired.

5.4 Continuous interpolations

We conclude with an evaluation of the continuous interpolations developed in Section 4. To this purpose we consider a convergence test based on a $10 \times 20\ mm^2$ specimen in plane strain with a pre-existing fully softened crack at its center and for half the width; see Figure 7. The Young modulus is 206.9 GPa

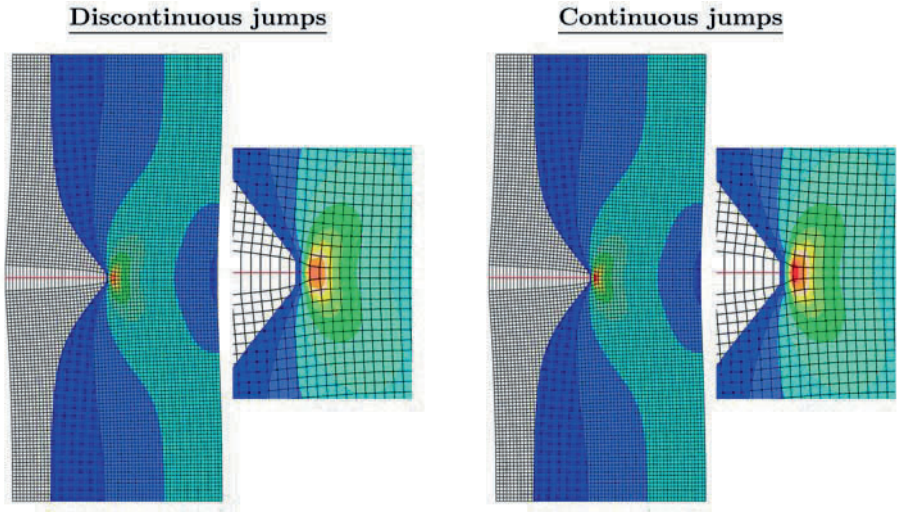


Fig. 7. Convergence test. Contours of the vertical stress with deformed meshes (scaled by 100) for the solutions with the discontinuous and the continuous interpolations of the jumps. The continuous interpolation also enforces a zero opening displacement at the crack tip, leading to a sharper stress concentration.

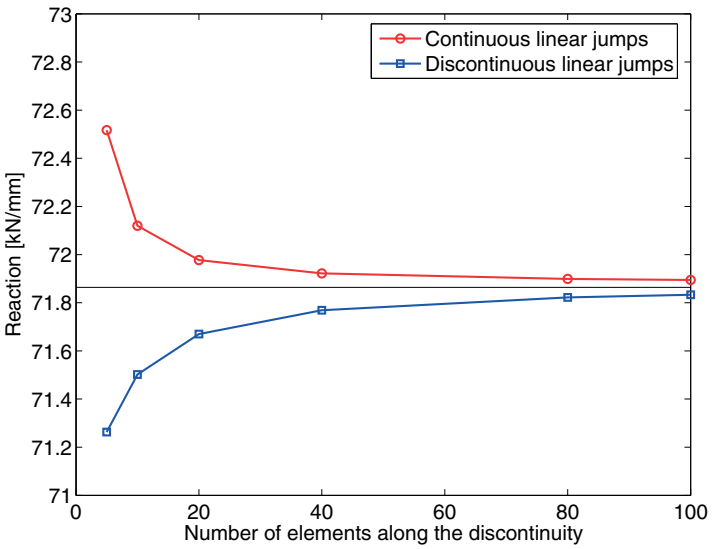


Fig. 8. Convergence test. Computed total reaction versus the number of segments resolving the crack for the different meshes considered.

and the Poisson ratio is 0.29. We compute the reaction obtained by imposing 1 mm displacement at the top for different structured meshes, and for that fixed length of the crack (i.e. the discontinuity does not propagate). Figure 7 depicts the solution for the mesh with a discontinuity resolved with 30 segments, showing the contours of the vertical stress for both discontinuous (piece-wise linear) and continuous interpolations of the jumps.

Figure 8 shows the convergence in the computed reaction versus the number of segments along the discontinuity for these two interpolations. A slightly better convergence (flatter curve) can be observed for the continuous interpolations. The stress contours in Figure 7 show also the sharper stress concentration for the continuous interpolations, noting that in this case we also impose the zero jump condition (31) at the crack tip. No special treatment of the tip element has been considered for capturing the stress concentration, in either case. Still, this sharper resolution indicates a promising performance for the continuous interpolation, especially as it is obtained again with very simple modifications of the original implementation.

6 Concluding remarks

The results presented here show the improvement gained by the consideration of linear displacement jumps combined with general quadrilateral elements, including basic displacement-based, mixed and enhanced elements. We obtain the lack of stress locking for a linearly separating discontinuity, and the generality gained by a fully linear separation (i.e. in both the normal and tangential components), allowing the representation of a number of modes without a spurious transfer of stresses. Furthermore, this generality allows the development of globally continuous interpolations of the jumps along the discontinuity. Remarkably, a continuous interpolation has been obtained without upsetting the very efficient local structure of the original elements, in which the enhanced parameters are condensed out at the element level.

Acknowledgements. Financial support for this research was provided by the ONR under grant no. N00014-00-1-0306 and AFOSR under grant no. FA9550-05-1-0117 with UC Berkeley. This support is gratefully acknowledged.

References

[1] Alfaite, J; Simone, A and Sluys, LJ (2003) International Journal of Solids and Structures 40:5799–5817
[2] Armero, F (1999) Mechanics of Cohesive-Frictional Materials 4:101–131
[3] Armero, F (2001) Computer Methods in Applied Mechanics and Engineering 191:181–213

[4] Armero, F and Ehrlich D (2006) Computer Methods in Applied Mechanics and Engineering 195:1283–1324
[5] Armero, F and Garikipati, K (1995) Recent advances in the analysis and numerical simulation of strain localization in inelastic solids. In: Proceedings of the 4^{th} Computational Plasticity Conference.
[6] Armero, F and Garikipati, K (1996) International Journal of Solids and Structures 33:2863–2885
[7] Belytschko, T and Black, T (1999) International Journal for Numerical Methods in Engineering 45:601-620
[8] Bolzon, G (2001) Computational Mechanics 27:463–473
[9] Borja, RI and Regueiro, RA (2001) Computer Methods in Applied Mechanics and Engineering 190:2555–2580
[10] Callari, C and Armero, F (2002) Computer Methods in Applied Mechanics and Engineering 191:4371–4400
[11] Ehrlich, D and Armero, F (2005) Computational Mechanics 35:237–264
[12] Larsson, R; Steinmann P and Runesson K (1998) Mechanics of Cohesive-Frictional Materials 4:171-194
[13] Linder, C and Armero, F (2006) submitted
[14] Manzoli, OL and Shing PB (2006) Computers and Structures 84:742–757
[15] Moës, N; Dolbow, J and Belytschko, T (1999) International Journal for Numerical Methods in Engineering 46:131–150
[16] Mosler, J and Meschke, G (2003) International Journal for Numerical Methods in Engineering 57:1553–1576
[17] (1996) Oliver, J International Journal for Numerical Methods in Engineering 39:3575–3623
[18] Oliver, J and Huespe, AE (2004) Computer Methods in Applied Mechanics and Engineering 193:2987–3014
[19] Reinhardt HW (1984) Heron 29(2):1–42
[20] Simo, JC; Oliver, J and Armero, F (1993) analysis of strong discontinuities induced by softening solutions in rate independent solids Computational Mechanics 12:277–296
[21] Simo, JC and Rifai, MS (1990) International Journal for Numerical Methods in Engineering 29:1595–1638
[22] Steinmann P (1999) Mechanics of Cohesive-Frictional Materials 4:133-152
[23] Wells, GN and Sluys, LJ (2001) International Journal for Numerical Methods in Engineering 50:2667-2682
[24] Wells, GN; Sluys, LJ and de Borst, R (2002) International Journal for Numerical Methods in Engineering 53:1235-1256

Evolving Material Discontinuities: Numerical Modeling by the Continuum Strong Discontinuity Approach (CSDA)

J. Oliver[1], A.E. Huespe[2], S. Blanco[1] and D.L. Linero[3]

[1]E.T.S. Enginyers de Camins, Canals i Ports, Technical University of Catalonia (UPC). Campus Nord, U.P.C., Mòdul C-1, Jordi Girona 1-3, 08034 Barcelona, Spain; xavier.oliver@upc.edu
[2]CIMEC, Intec/Conicet, Guemes 3451, 3000 Santa Fe, Argentina
[3]Universidad Nacional de Colombia, Colombia

Summary. The CSDA, as a numerical tool for modeling evolving displacement discontinuities in material failure problems, is addressed. Its specific features are: a) the explicit use of a (regularized) strong discontinuity kinematics, b) the introduction of the material failure constitutive model in a continuum (stress-strain) format, and c) the determination of the onset and propagation of the discontinuity by means of constitutive model material bifurcation analysis. Numerical applications to concrete failure and soil collapse problems are presented.

Key words: Continuum Strong Discontinuity Approach, material failure simulation, IMPL-EX algorithm.

1 Introduction

Displacement discontinuities are observed, at macroscopic scales, associated to material failure in solids, as cracks, fractures, shear bands, etc. Therefore, modeling its formation, evolution and propagation is a specific goal of Computational Material Failure Mechanics.

In material science, numerical simulation of problems involving quasi-brittle or ductile fracture, dynamic fracture propagation, fragmentation or spalling, does not only serves as a way of obtaining the response of a given model, but also as a tool for a better understanding of theoretical, and not sufficiently well known, aspects of those phenomena. This is, typically, the case of dynamic fracture problems, where new analytical developments are extremely difficult to develop, and experimental results can be then approached by numerical procedures so that the interaction between different mechanisms can be better understood and explained.

Alain Combescure et al. (eds.), IUTAM Symposium on Discretization Methods for Evolving Discontinuities, 123–138.

At the beginning of 90s the cohesive model approach showed fairly good performance for reaching that objective [1]. It is now well established that the adoption of cohesive models requires, in addition to the standard continuum model for the bulk, three fundamental ingredients:

1. a traction-separation law for describing the loss of cohesion in the fracture process zone;
2. a crack growth criteria for determining the inception time of the discrete law and the propagation trajectory of the discontinuity across the solid.

To reproduce the fracture formation and its subsequent evolution, both ingredients must interact with an enriched kinematics model allowing jumps in the velocity field, from now on termed *strong discontinuities* [18]. Thus, a third ingredient is necessary:

3. an enhanced kinematics including strong discontinuity modes.

In recent years an alternative setting for displacement discontinuity modeling, involving similar mechanical and numerical ingredients, has appeared. This approach, which is closely related to the cohesive one, and termed *strong discontinuity approach* (SDA), exhibits two specific branches:

(i) the *Discrete Strong Discontinuity Approach* (DSDA), which is similar to the cohesive model in terms of the given treatment of the traction-separation law and the crack growth criteria [1, 2, 5, 17].
(ii) the *Continuum Strong Discontinuity Approach* (CSDA) where continuum (stress-strain) constitutive models, equipped with strain softening, are used to model the material response in the fracture interface.

In this work, the last approach is addressed. Its main ingredients are presented and analyzed and its capability evaluated through two typical problems of material failure simulation.

The remaining of the work is organized as follows: in Section 2, the strong discontinuity kinematics is presented. Section 3 describes the fundamental topic of the CSDA i.e. the treatment given to the cohesive forces in the sense of the above mentioned ingredients. This aspect is what confers to the CSDA a specific character, making it essentially different from the classical cohesive and DSDA models.

The variational formulation of the problem is presented in Section 4, in a general way such that different available finite element techniques, like EFEM or XFEM methods, can be straightforwardly inserted. Then, in Section 5, a specific procedure for integration of the softening constitutive model, the IMPL-EX method, which provides high robustness and stability to the numerical simulations, is tackled. In Section 6 some representative examples are presented and, finally, in Section 7 conclusions about the work are supplied.

2 Strong Discontinuity Kinematics

Let Ω be an open body. The strong discontinuity kinematics is defined in Ω by considering the existence of displacement jumps, $\boldsymbol{\beta}$, in the displacement field $\mathbf{u}$, across a surface S, as follows:

$$\mathbf{u}(\mathbf{x}) = \bar{\mathbf{u}}(\mathbf{x}) + \mathcal{H}_S \boldsymbol{\beta}(\mathbf{x}); \quad \mathcal{H}_S = \begin{cases} 1 \ \forall \mathbf{x} \in \Omega^+ \\ 0 \ \forall \mathbf{x} \in \Omega^- \end{cases} \tag{1}$$

where $\bar{\mathbf{u}}$ represents a smooth displacement field in Ω and $\mathcal{H}_S$ the Heaviside's step function. The discontinuity surface S divides Ω in two disjoint parts Ω^+ and Ω^-, see Figure 1.

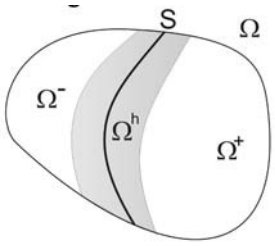


Fig. 1. Strong discontinuity in a body.

In [19], a slightly different functional representation of the discontinuous displacement $\mathbf{u}$, more adapted to a finite element approach, has been defined as follows: let us define a region $\Omega^h \subset \Omega$ enclosing S, $(S \subset \Omega^h)$, and a smooth function φ^s, such that: $\varphi^s = 1 \ \forall \mathbf{x} \in \Omega^+ / \Omega^h$ and $\varphi^s = 0 \ \forall \mathbf{x} \in \Omega^- / \Omega^h$. Then, the discontinuous displacement field $\mathbf{u}$ can be alternatively written as:

$$\mathbf{u}(\mathbf{x}) = \bar{\mathbf{u}}(\mathbf{x}) + \mathcal{M}_S \boldsymbol{\beta}(\mathbf{x}); \quad \mathcal{M}_S = \mathcal{H}_S - \varphi^s, \tag{2}$$

where $\mathcal{M}_S$ is the so called unit jump function with support Ω^h (see Figure 1). This allows that, if Ω^h is a small region not including the body boundary where Dirichlet's conditions are imposed, the displacement on that boundary is imposed on both $\mathbf{u}$ and $\bar{\mathbf{u}}$ fields.

A key aspect of the CSDA is to assume a generalization of those kinematical concepts on S, like strains and derived magnitudes, which in the classical continuum approach lose their bounded character due to the displacement jump, by introducing the generalized gradient concept. Then, the strains being compatible with the displacement jump (2), are defined as:

$$\boldsymbol{\varepsilon}(\mathbf{x}) = \nabla^{\text{sym}} \mathbf{u}(\mathbf{x}) = \underbrace{\bar{\boldsymbol{\varepsilon}}(\mathbf{x})}_{\substack{\text{regular} \\ \text{(bounded)}}} + \underbrace{\delta_S (\boldsymbol{\beta} \otimes \mathbf{n})^{\text{sym}}}_{\substack{\text{singular} \\ \text{(unbounded)}}}, \tag{3}$$

where a singular term appears, affected by the Delta's function δ_S coming from the displacement strong discontinuity mode. The structure of the singular term in (3) deserves a special attention. It is the tensorial product between the jump displacement vector $\boldsymbol{\beta}$ and the unit normal vector $\mathbf{n}$ to the discontinuity surface S and pointing to Ω^+.

3 Continuum Constitutive Models and Traction-Separation Laws

In the context of the classical cohesive models, the continuum behavior is described by means of a constitutive model that is completely independent from the traction-separation law acting at the fracture process zone. In support of this methodology, it could be argued that the physical phenomena taking place in the material at the fracture process zone, are essentially different from those governing the continuum deformation mechanisms. However, for numerical simulation purposes, two key aspects of this methodology should be questioned:

1. it requires the definition of a specific constitutive law for the cohesive forces as a new ingredient with respect to the continuum constitutive model at the bulk. Therefore, its characterization requires additional laboratory tests;
2. The transition from the continuum to the discrete constitutive law, in a given material point, cannot be arbitrarily done. Requirements as the so called time continuity [8] impose specific constraints not always fulfilled in existing cohesive models.

Actually, for numerical simulation purposes it is very common to use ad-hoc cohesive laws, similar to damage or plastic models, in the reduced space of tractions versus displacement jumps. These cohesive laws are not necessarily derived from a micromechanical models and, therefore, they can not capture or predict the basic phenomenology arising in the fracture process zone (see for instance [8, 9, 16]).

As an alternative to this approach, the CSDA proposes to obtain the cohesive law consistently from a continuum constitutive model, by means of its projection onto the discontinuity interface S (see Figure 1). For this purpose, a continuum constitutive law, equipped with strain softening, is assigned to the material in any point, in such a way that a material bifurcation process and the strong discontinuity kinematics impose the *natural* fulfillment of that projected traction-separation law, at the discontinuity interface, as the discontinuity onsets. In this way the passage from the continuum to the discrete law is a smooth mapping in time. Of course, this procedure requires a continuum constitutive law, rich enough for capturing the physical phenomena appearing in the strain localized zone. Then, once defined the continuum model, the co-

hesive force is derived automatically as its projection into the failure interface [10]. More details can be found in [11] and [12].

3.1 Damage and Plasticity Continuum Models

Let us consider the two classical continuum material models of Table 1: (i) a continuum isotropic damage model and (ii) a J_2 elasto-plastic model.

The notation used in Table 1 is the following: for the plastic model, the total strain $\boldsymbol{\varepsilon}$ is partitioned in the elastic part $\boldsymbol{\varepsilon}^e$, and the plastic counterpart $\boldsymbol{\varepsilon}^p$. The stress tensor is denoted by $\boldsymbol{\sigma}$ and its deviatoric part by $\boldsymbol{\sigma}^{\text{dev}}$, $\mathbf{C}^e$ is the Hooke's elastic tensor, r and q represent, respectively, the strain and stress-like internal variables of the model, and ϕ_ε and $\boldsymbol{\varepsilon}_\sigma$ define the elastic domain in each case. Identical notation, $\boldsymbol{\sigma}_u$, has been used for the ultimate stress in the damage model and the yield stress in the plastic one. The Young's modulus is denoted by E and H is the softening modulus. In the damage model $\bar{\boldsymbol{\sigma}}^+$ is the positive part of the effective stress tensor $\bar{\boldsymbol{\sigma}}$, defined as

$$\bar{\boldsymbol{\sigma}}^+ = \sum_1^3 \langle \bar{\sigma}_i \rangle \mathbf{p}_i \otimes \mathbf{p}_i,$$

where $\langle \bar{\sigma}_i \rangle$ is the McAulay bracket of the ith principal stress, $\bar{\sigma}_i$, and $\mathbf{p}_i$ stands for the ith eigenvector of $\bar{\boldsymbol{\sigma}}$.

Table 1. Continuum damage and plastic material models.

Isotropic continuum damage model with strain softening	J_2 elasto-plastic model with isotropic strain softening	
$\boldsymbol{\sigma} = \frac{q}{r} \underbrace{\mathbf{C}^e : \boldsymbol{\varepsilon}}_{\bar{\boldsymbol{\sigma}}} = \frac{q}{r}\bar{\boldsymbol{\sigma}}$	$\boldsymbol{\sigma} = \mathbf{C}^e : \varepsilon^e = \mathbf{C}^e : (\varepsilon - \varepsilon^p)$ $\varepsilon = \varepsilon^e + \varepsilon^p$	(4)
$\dot{r} = \dot{\lambda};\ r\|_{t=0} = r_0 = \sigma_u/\sqrt{E}$	$\dot{r} = \dot{\lambda}; \quad r\|_{t=0} = 0$ $\dot{\boldsymbol{\varepsilon}}^p = \dot{\lambda}\boldsymbol{\xi}; \quad \boldsymbol{\xi} = \partial_\sigma \phi_\sigma$	(5)
$\phi_\varepsilon(\boldsymbol{\varepsilon}, r) \equiv \tau_{\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}) - r$ $\tau_\varepsilon(\boldsymbol{\varepsilon}) \equiv \sqrt{\bar{\boldsymbol{\sigma}}^+ \cdot \mathbf{C}^{e^{-1}} \cdot \bar{\boldsymbol{\sigma}}}$	$\phi_\sigma(\boldsymbol{\sigma}, q) = J_2(\boldsymbol{\sigma}) - (\sigma_u - q)$ $J_2(\boldsymbol{\sigma}) = \sqrt{3/2}\|\boldsymbol{\sigma}^{\text{dev}}\|$	(6)
$\dot{\lambda} \geq 0;\ \phi_\varepsilon \leq 0;\ \dot{\lambda}\phi_\varepsilon = 0$	$\dot{\lambda} \geq 0;\ \phi_\sigma \leq 0;\ \dot{\lambda}\phi_\sigma = 0$	(7)
$\dot{q} = H\dot{r}; \quad q \geq 0$ $q\|_{t=0} = r_0 = \sigma_u/\sqrt{E}; \quad q\|_{t=\infty} = 0$	$\dot{q} = -H\dot{r}; \quad q \geq 0$ $q\|_{t=0}; \quad q\|_{t=\infty} = \sigma_u$	(8)

Projected damage and plasticity models

The projection of both models onto the discontinuity surface has been presented in [10]. There, it has been shown that imposing the traction continuity condition across S, it is possible to derive the functional relation between $\mathbf{t}_S = \boldsymbol{\sigma}_S \cdot \mathbf{n}$, the traction vector on S, and the displacement jump vector $\boldsymbol{\beta}$. This becomes the discrete traction-separation law of the associated (projected) cohesive model. Applying this procedure to the classical models (4)–(8), it determines the traction-separation laws described in Table 2. In this table, λ and μ denote the Lame's parameters, q and α internal variables, while the (intrinsic) softening modulus $\bar{H}_S$ comes from the regularization of the continuum softening modulus H.

Table 2. Discrete damage and plastic models.

Isotropic continuum damage model with strain softening	J_2 elasto-plastic model with isotropic strain softening	
$\mathbf{t}_S = (1-\omega)\mathbf{Q}^e \cdot \boldsymbol{\beta}; \quad \omega = \dfrac{q}{\alpha}$ $\mathbf{Q}^e = \mathbf{n} \cdot \mathbf{C}^e \cdot \mathbf{n} = (\lambda+\mu)\mathbf{n} \otimes \mathbf{n} + \mu \mathbf{1}$	$\dot{\boldsymbol{\beta}} = \dot{\gamma}\mathbf{m}_S; \qquad \dot{\boldsymbol{\beta}} \cdot \mathbf{n} = 0$ $\mathbf{m}_S = \sqrt{\dfrac{3}{2}} \dfrac{\mathbf{t}_S^{\text{dev}}}{\|\mathbf{t}_S^{\text{dev}}\|}; \quad \mathbf{t}_S^{\text{dev}} = \boldsymbol{\sigma}_S^{\text{dev}} \cdot \mathbf{n}$	(9)
$\dot{q} = \bar{H}_S \dot{\alpha}; \quad \dot{\alpha} = \dot{\gamma}; \alpha(t)\|_{t=t_d} = 0$ $q(t)\|_{t=t_d} = q_d = \dfrac{\sigma_u}{\sqrt{E}} > 0$	$\dot{q} = \bar{H}_S \dot{\alpha}; \quad \dot{\alpha} = \dot{\gamma}$ $q(0) = 0$	(10)
$\mathsf{F}_t(\mathbf{t}, q) = \tau_t - q$ $\tau_t = \underbrace{\sqrt{\mathbf{t}_S \cdot (\mathbf{Q}^e)^{-1} \cdot \mathbf{t}_S}}_{\text{effective traction}}$	$\phi_S(\mathbf{t}_S^{\text{dev}}, q_s) =$ $= \sqrt{3(\mathbf{t}_S^{\text{dev}} \cdot \mathbf{t}_S^{\text{dev}})} - (\sigma_u - q)$	(11)
$\hat{\gamma} \geq 0; \ \mathsf{F}_t(\mathbf{t}, q) \leq 0; \ \hat{\gamma}\mathsf{F}_t = 0$	$\hat{\gamma} \geq 0; \ \phi_S \leq 0; \ \hat{\gamma}\phi_S = 0$	(12)

In a loading process leading to the complete material exhaustion, it is possible to characterize $\bar{H}_S$ by requiring that the dissipated energy in the traction-separation law be equal to the fracture energy G_f (see [11]).

Strong discontinuity activation criteria

This ingredient is introduced into the CSDA by resorting to the material instability condition, derived from an analytical analysis of the bifurcation condition in the loading process. From the theoretical point of view, a material exhibiting strain softening is prone to bifurcate exhibiting a subsequent unstable regime. In a loading process this determines the possibility of developing an alternative mode of deformation, at the material point $\mathbf{x}$ and t at the bifurcation time t_B, as the following condition is fulfilled:

$$\det(\mathbf{Q}^{\text{loc}}(\mathbf{x}, \mathbf{n}, t_B)) = 0,$$
$$\mathbf{Q}^{\text{loc}}(\mathbf{x}, \mathbf{n}, t_B) = \mathbf{n} \cdot \mathbf{C}^{\text{load}} \cdot \mathbf{n}, \tag{13}$$

where $\mathbf{C}^{\text{load}}$ is the continuum constitutive tangent tensor in loading stage. Criterion (13) indicates that a bifurcation state is reached, the first time, t_B, that the acoustic tensor $\mathbf{Q}^{\text{loc}}$ becomes singular for some direction $\mathbf{n}$. This condition, which is necessary but not sufficient to induce a strong discontinuity deformation mode, depends on the material stress-strain state.

4 Variational Formulation and Finite Element Approach

A straightforward way to capture strong discontinuities across the element boundaries, in a finite element mesh, consists of adding, for every regular node of the mesh, additional degrees of freedom, which are freed when the criteria of fracture propagation is reached in any of the neighbor elements. The traction forces are implemented via interface finite elements. This methodology, which enforces the discontinuity path to follow the element boundaries, has been widely used in the context of cohesive model approaches [16].

An alternative approach consists of embedding the strong discontinuities into the finite elements, as proposed in the so called *finite elements with embedded discontinuities* (EFEM) or in the *extended finite element techniques* (XFEM). Both of them can be identically formulated, into the same variational setting, by choosing adequate test function spaces.

Let the problem governing equations be written as a classical variational (Virtual Work) principle:

$$\text{Find: } \mathbf{u}^h \in \mathcal{V}^h; \quad \text{such that}$$
$$\int_{\Omega^h} \nabla^{\text{sym}} \delta\mathbf{u}^h \cdot \sigma d\Omega = \int_{\Gamma_\sigma^h} \delta\mathbf{u}^h \cdot \bar{\mathbf{t}} d\Gamma_\sigma; \ \forall \delta\mathbf{u}^h \in \mathcal{V}_0^h, \tag{14}$$

where, for the XFEM technology, the discrete functional space $\mathcal{V}^h$ is given by a discontinuous displacement field, see also Belytschko et al. [4]:

$$\mathcal{V}_{\text{XFEM}}^h = \left\{ \mathbf{u}^h(\mathbf{x}) \mid \mathbf{u}^h(\mathbf{x}) = \sum_{i=1}^{n_{\text{node}}} (N_i(\mathbf{x})\mathbf{d}_i + \mathcal{H}_S N_i(\mathbf{x})\boldsymbol{\beta}_i) \right\}, \tag{15}$$

N_i being the standard shape function of the basic finite element, triangle, quadrilateral, etc.; $\mathbf{d}_i$ the interpolation parameters corresponding to the smooth displacement field and $\boldsymbol{\beta}_i$ the nodal parameters of the strong discontinuity field (actually $\boldsymbol{\beta}_i$ represent displacement jumps at node i). Alternatively, for the EFEM technology, the discrete functional space is defined by:

$$\mathcal{V}^h_{\text{EFEM}} = \left\{ \mathbf{u}^h(\mathbf{x}) \mid \mathbf{u}^h(\mathbf{x}) = \sum_{i=1}^{n_{\text{node}}} (N_i(\mathbf{x})\mathbf{d}_i) + \sum_{e=1}^{n_{\text{elem}}} \mathcal{M}_S^{(e)} \boldsymbol{\beta}_e \right\},$$

$$\mathcal{M}_S^{(e)} = \mathcal{H}_S - \varphi^{(e)}; \quad \varphi^{(e)} = \sum_{i=1}^{n^{(e)}_{\text{node}^+}} N_i^{(e)}, \tag{16}$$

with $\mathcal{M}_S^{(e)}$ being a discontinuous enriching function, the discrete unit jump function, with support on element e, $\boldsymbol{\beta}_e$ are the parameters corresponding to the elemental displacement jump and $n^e_{n^+_{\text{node}}}$ are those elemental nodes belonging to Ω^+.

For this second case, the variations respect to each parameter $(\mathbf{d}_i, \boldsymbol{\beta}_e)$ results in the following set of equations:

$$\delta\mathbf{d}_i \cdot \left(\int_{\Omega^h} \nabla N_i \cdot \boldsymbol{\sigma} d\Omega - \int_{\Omega^h_\sigma} N_i \cdot \bar{\mathbf{t}} d\Gamma \right) = 0 \quad \forall \delta\mathbf{d}_i; \quad i = 1, n_{\text{node}},$$

$$\delta\boldsymbol{\beta}_e \cdot \left(\int_{\Omega^{(e)}} \nabla \varphi^{(e)} \cdot \boldsymbol{\sigma} d\Omega + \int_{S^{(e)}} (\boldsymbol{\sigma}_S \cdot \mathbf{n}) dS \right) = 0 \quad \forall \delta\boldsymbol{\beta}_e; \quad e = 1, n_{\text{elem}}. \tag{17}$$

This is the EFEM discrete weak form of the problem governing equations. The variational problem (14), using either the spaces (15) or (16) determines a symmetric finite element approach.

A comparative analysis between both techniques EFEM vs. XFEM has been presented in Oliver et al. [14].

4.1 Strong Discontinuity Tracking Algorithm

Both the EFEM and XFEM technologies, require an additional ingredient: the determination of the discontinuity path and its intersection with the mesh elements. Different algorithms have been developed to reach this objective: (i) *global tracking algorithms*, where the discontinuity paths are determined *globally* (for the complete mesh at once) as presented in Oliver et al. [13], or the *level set method* of Gravouil et al. [7], Feist and Hofstetter [6], etc; and (ii) local tracking algorithms, where the path is determined in an element to element way.

From the authors' experience, global tracking algorithms are more suitable for 3D analyzes. This type of algorithm can be described as follows:

Let us assume that a vector field $\mathbf{n}(\mathbf{x})$, orthogonal to the discontinuity path, is defined at every point (Gauss Point in the finite element mesh) of the solid by means of the criterion of Section 3.1.2. In that part of the body where an elastic response is observed, an admissible direction $\mathbf{n}$ is obtained by assuming an, artificially reduced, value for the yield stress σ_u such that the stress-strain state just corresponds with the loading case for the chosen non-linear model.

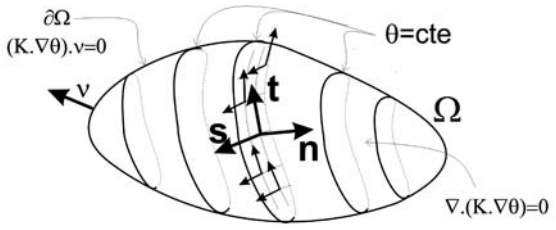


Fig. 2. Thermal-like tracking algorithm.

In the plane orthogonal to $\mathbf{n}$, two, again orthogonal, unit vectors, $\mathbf{t}(\mathbf{x})$ $\mathbf{s}(\mathbf{x})$, are considered so that they define, at point $\mathbf{x}$, the tangent plane of propagation of the discontinuity (see Figure 2). Then, the scalar field $\theta(\mathbf{x})$ is the solution of the following problem:

$$\text{Find: } \theta \text{ such that}$$

$$\nabla \cdot \mathbf{K}_\theta \nabla\theta = 0; \quad \mathbf{K}_\theta(\mathbf{x}) = \mathbf{t} \otimes \mathbf{t} + \mathbf{s} \otimes \mathbf{s} \tag{18}$$

defines the family of level sets $S_c \equiv \{\mathbf{x} \mid \theta(\mathbf{x}) = c; \mathbf{x} \in \Omega\}$ $\forall c \in R$, which are the envelopes of the vector fields $\mathbf{t}(\mathbf{x})$ and $\mathbf{s}(\mathbf{x})$ (see [13] for additional details).

From the numerical point of view, the accompanying problem (18), equivalent to a linear adiabatic thermal one with a non-homogenous anisotropic conductivity $\mathbf{K}_\theta$, can be solved when every, or once several, time step has converged. Once obtained the field θ, managing the discontinuity evolution across the body is rather straightforward, considering the following aspects:

(i) Once the first element in the mesh reach the critical condition (13), the iso-level surface crossing that element determines the first crack propagation path "θ_1";
(ii) Elements subsequently reaching the critical condition, which are not crossed by the previously defined iso-level surfaces θ_j $(j = 1, 2, \ldots, i)$, define new crack pathes θ_{i+1}, θ_{i+2}, ...
(iii) Once an element has reached the critical condition, the normal vector $\mathbf{n}$ and the θ field's d.o.f. in that element are frozen. While the element that is intersected by the θ_i path does not reach the critical condition, the θ field could be modified in that point.
(iv) In the finite element mesh, is possible to define a shield zone where some elements are precluded to reach the critical condition.

When using an implicit procedure, the item (iv) becomes a very important issue to get an adequate algorithmic robustness. Nevertheless, the use of the IMPL-EX algorithm, to be described in the next section, has relaxed this restriction.

It is worth mentioning here that some new methods of embedded strong discontinuities do not require a tracking algorithm [17].

5 Constitutive Equation Integration. IMPL-EX Scheme

In computational material failure mechanics, the constitutive equation integration scheme deserves an important attention considering the large computational time savings that can be obtained with a correct choice. Historically, two widely acknowledged algorithms have been used: (i) *explicit* and (ii) *implicit* algorithms. Explicit algorithms, do not require neither evaluating the stiffness matrix nor inverting it, but demand a much larger number of time increments in comparison with the implicit ones (one or two orders larger). On the other hand, when using implicit algorithms in computational material failure mechanics, numerical simulations exhibit a dramatic loss of robustness, so that their most important advantage (larger time steps), gets reduced and, many times, they become not competitive.

The authors have obtained remarkable good results by using a new implicit-explicit (IMPL-EX) integration technique (see [15]). It consists of a double integration of the constitutive model at every time step n. The stresses and internal variables are determined by: (i) a standard *implicit* Euler-Backward scheme, resulting in implicit values of the stresses $\boldsymbol{\sigma}_n^*$ and the internal variables set r_n^*, and (ii) an *explicit extrapolation of the internal variable set*, $\tilde{r}_n$, in terms of the implicit values at previous time steps, from which the corresponding stresses, $\tilde{\boldsymbol{\sigma}}_n = \boldsymbol{\sigma}(\tilde{r}_n, \boldsymbol{\varepsilon}_n)$ are computed and used for determining the corresponding internal forces. Then, it can be proven that the corresponding consistent tangent algorithm $\mathbf{C}_n^{\text{eff}} = \partial\tilde{\boldsymbol{\sigma}}_n/\partial\boldsymbol{\varepsilon}_n$ becomes strictly positive definite, which has striking effects on the robustness of the numerical response. Moreover, for a larger number of the, most commonly used, constitutive models, $\mathbf{C}^{\text{eff}}$ becomes constant and, therefore, the linearized system of equations takes just one iteration to converge at every time step. The result is a dramatic reduction of the resulting computational effort at the cost of introducing an additional (time step length dependent) error.

We present an overview of this procedure, by assuming a J_2 plasticity model. Let the implicit stresses, $\boldsymbol{\sigma}^*$ and internal variables r^*, q^*, at previous time steps, be known. The IMPL-EX values of the stresses, and internal variables, at step n are determined by the following algorithm:

$$\Delta\tilde{r}_n = \Delta\tilde{\gamma}_n = \frac{\Delta r_{n-1}^*}{\Delta t_{n-1}}\Delta t_n = (r_{n-1}^* - r_{n-2}^*)\frac{\Delta t_n}{\Delta t_{n-1}}; \quad \tilde{q}_n = q_{n-1}^* + H\Delta\tilde{r}_n,$$

$$\tilde{\boldsymbol{\sigma}}_n = \boldsymbol{\sigma}_n^* + \mathbf{C}^e : (\Delta\boldsymbol{\varepsilon}_n - \Delta\tilde{\boldsymbol{\varepsilon}}^p); \quad \Delta\tilde{\boldsymbol{\varepsilon}}^p = \Delta\tilde{\gamma}_n \frac{\tilde{\mathbf{s}}_n}{\|\hat{\mathbf{s}}_n\|},$$

$$\tilde{\mathbf{s}}_n = \tilde{\boldsymbol{\sigma}}_n^{\text{dev}} = \mathbf{I}^{\text{dev}} : \tilde{\boldsymbol{\sigma}}_n; \quad \|\hat{\mathbf{s}}_n\| - \sqrt{\frac{2}{3}}(\sigma_Y - \tilde{q}_n) = 0, \tag{19}$$

where $\mathbf{I}^{\text{dev}}$ is the fourth-order identity deviatoric tensor. It is straightforward to compute the corresponding algorithmic tangent tensor as:

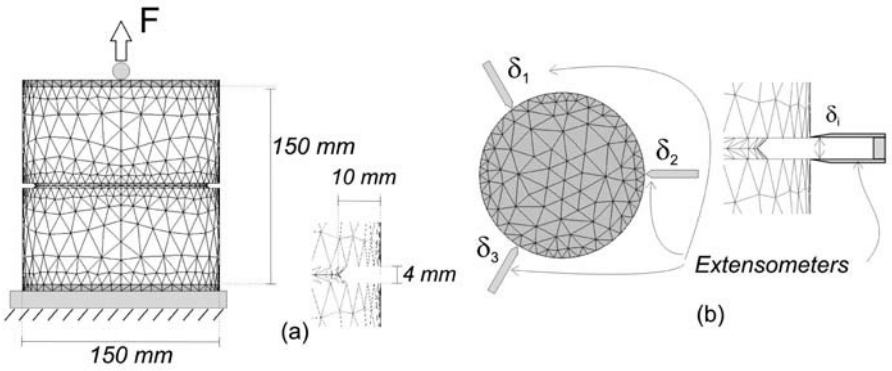


Fig. 3. Notched cylindrical bar. Finite element mesh.

$$\mathbf{C}_n^{\text{eff}} = \left[\mathbf{C}^{e^{-1}} + \sqrt{\frac{3}{2}} \frac{\Delta\tilde{\gamma}_n}{(\sigma_Y - \tilde{q}_n)} \mathbf{I}^{\text{dev}} \right]^{-1}, \tag{20}$$

which is positive definite and constant (during the time step n) since $\Delta\tilde{\gamma}_n$ and $\tilde{q}_n$ are known at the beginning of the time step.

6 Numerical Applications

In this section two problems are presented, which show the ability of the CSDA for capturing typical material failure mechanisms. The first one corresponds to a concrete specimen undergoing a cracking process, what produces the complete loss of its structural strength. The second application refers to the formation of a slip surface. In this case, the continuum material constitutive law follows a J_2 plasticity model.

Both solutions have been obtained by means of 3D finite element simulations and the discontinuity surfaces were determined by using the tracking algorithm described in Section 4.1.

6.1 Notched Cylindrical Bar in Tensile State

This test has been reported by Barragan [3], from where the experimental results, to validate the present numerical model, have been taken. It consists of a notched cylindrical bar under uniaxial loading. The top of the bar, where the force is imposed, is free to rotate.

The specimen geometrical dimensions are displayed in Figure 3. The parameters of the material, simulated by means of the damage model of Section 3.1, are: Young's modulus $E = 30.5$ GPa, Poisson's ratio: $\nu = 0.2$, ultimate stress $\sigma_u = 1.79$ MPa and fracture energy $G_f = 100$ N/m.

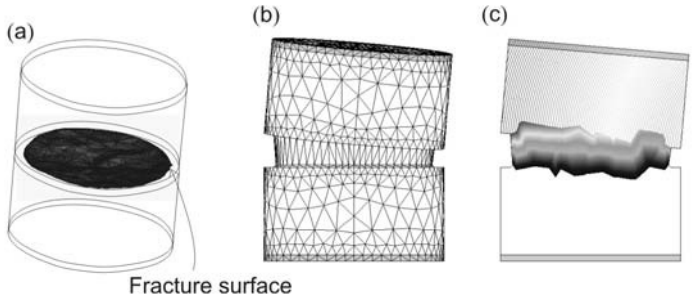


Fig. 4. Notched cylindrical bar. Deformation.

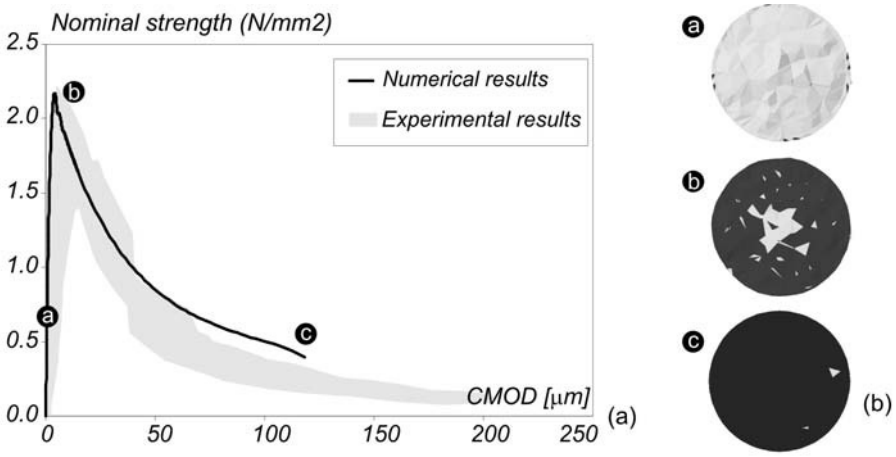


Fig. 5. Notched cylindrical bar. Results.

The finite element mesh consists of 8543 tetrahedra, as displayed in the same figure. It is important to remark that the elements in the neighborhood of the notched zone have been arbitrarily generated so that they are not constrained to lying between the two horizontal planes defining the horizontal notch sides. It is expected that the fracture surfaces define a quasi-planar surface crossing the notched section. In consequence, this mesh design is a challenge for the correct capturing of the strong discontinuity surface when tetrahedral finite elements are used.

Figure 4 displays the discontinuity interface determined by the tracking algorithm, the deformed mesh and the surfaces of iso-displacement at the end of analysis. It can be observed the rotation of the upper section. This rotation determines a non-uniform circumferential CMOD. Therefore, to evaluate this magnitude, three extensometers, distributed as shown in Figure 3b, were used in the test.

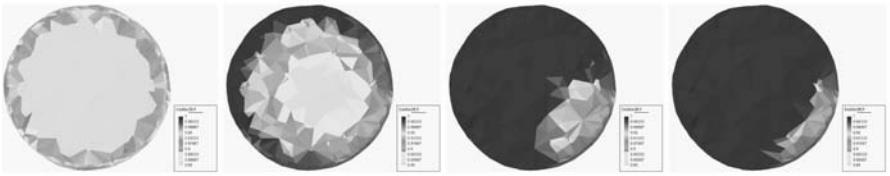


Fig. 6. Notched cylindrical bar: damage evolution of the notched section.

Figure 5a plots the nominal strength vs. the CMOD curve which is obtained as the mean value of the three extensometers. The nominal strength σ_n is defined as:

$$\sigma_n = \frac{4P_u}{\pi D^2}, \tag{21}$$

where P_u is the collapse load and $D = 130$ mm the notched section diameter. The same figure displays the experimental results. It can be observed the god fitting of the numerical solution mainly in the post critical regime.

The load increase determines a strain localization process in the notch zone, whose evolution is shown in Figure 5b. The finite elements whose discontinuity mode is active are represented in black. Due to the rotation of the specimen half upper part, with respect to the lower part, the progress of the discontinuity surface does not shown axi-symmetry. This result is checked in Figure 6, where the damage evolution is shown in the same bar section.

6.2 Embankment Stability Analysis

The stability analysis of a 3D embankment, constituted by a deformable soil lying over almost rigid bedrock, is performed. Its geometry, boundary conditions and finite element mesh are displayed in Figure 7. Notice the bedrock step influencing the structural collapse mechanism.

The material model assumed for the soil is a J_2 plasticity one, whereas the rock is assumed elastic. Their material parameters are: (i) *deformable soil*: Young's modulus $E = 10$ MPa, Poisson's ratio: $\nu = 0.4$; yield stress $\sigma_y = 0.1$ MPa and fracture energy $G_f = 10$ kN/m, density $\rho = 1800$ kg/m^3; (ii) *rigid bedrock*: $E = 100$ MPa, Poisson's ratio: $\nu = 0.4$, density $\rho = 1800$ kg/m^3.

The external action is the self weight, and the load factor is applied on the material density. For the numerical solution strategy an arc-length method, controlling the vertical displacement at point P, is used.

The finite element mesh consists of 5250 bricks using the BBAR technology to avoid the kinematics locking induced by the isochoric plastic deformation.

Figure 8 displays the slip surface captured by the numerical simulation. It is observed the two curved segments resulting from the bedrock step.

Figure 9a shows the structural response by means of the load factor vs. the vertical displacement at point P. The critical load corresponds to a density $\rho = 4896$ kg/m^3 (load factor 2.72).

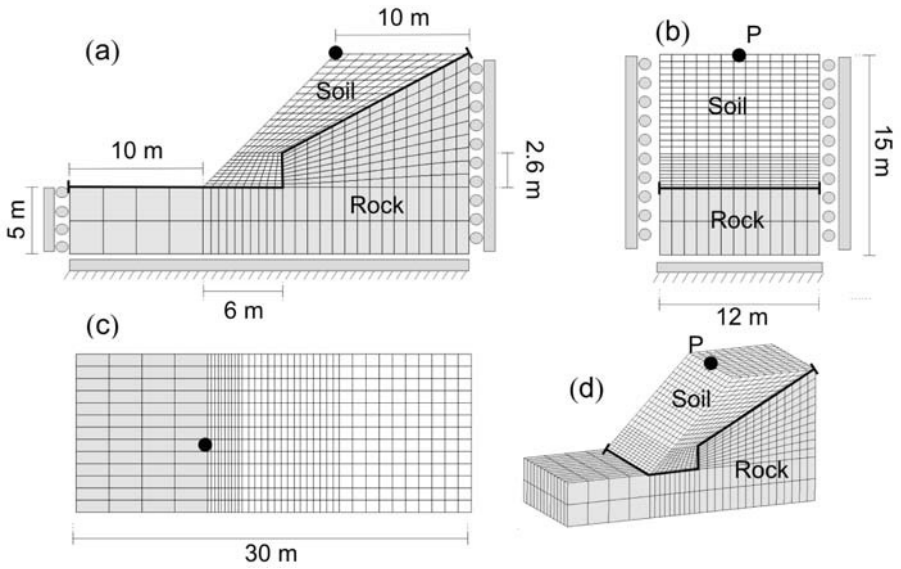


Fig. 7. Embankment stability analysis: geometry and finite element model.

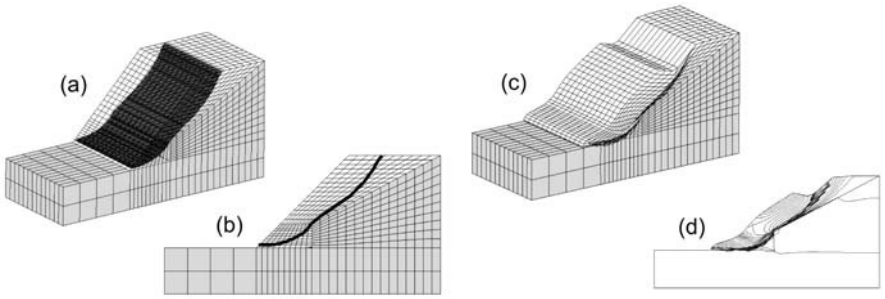


Fig. 8. Embankment stability analysis: slip surface, deformed mesh and isosurfaces of displacements.

In correspondence, Figure 9b shows, in black, the sequence of the strong discontinuity activation along the captured slip surface.

7 Conclusions

The CSDA, in conjunction with embedded finite element (EFEM) techniques and the implicit-explicit (IMPL-EX) time integration scheme, constitutes a robust approach to the numerical simulation of material failure. It allows general 3D material failure simulations in those problems where evolving displacement discontinuities characterize the collapse mechanism. The most specific

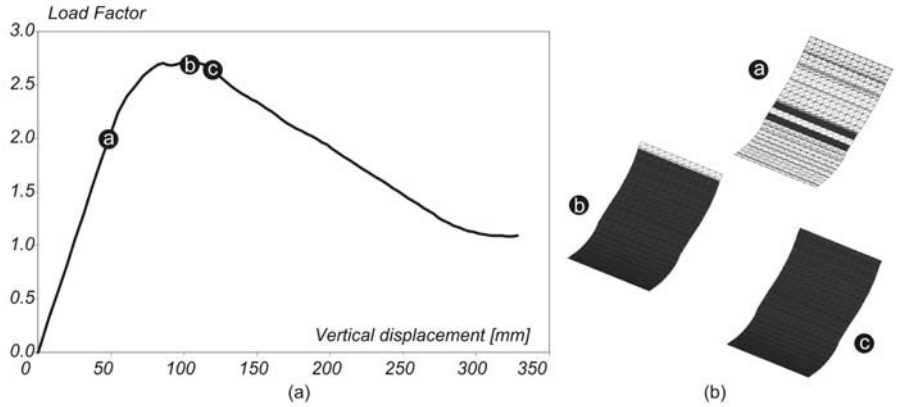


Fig. 9. Embankment stability analysis: load factor vs. vertical displacement at point P; evolution along the time of the sliding surface.

features of the CSDA in front of alternative procedures are: (a) the (regularized) strong discontinuity kinematics is explicitly introduced in the strain description, (b) specific traction-separation laws are not required and the failure constitutive model is introduced in a continuum (stress-strain) format and (c) the onset and propagation of the discontinuity is determined via a material bifurcation analysis of the continuum tangent constitutive operator.

In this work, an overview of those ingredients has been provided, and the method performance has been illustrated by its application to representative examples.

Acknowledgements

Financial support from the Spanish Ministry of Science and Technology trough grants BIA2005-09250-C03-03 and BIA2004-02080, is gratefully acknowledged.

References

1. Alfaiate J., Sluys L.J. (2004). Discontinuous numerical modelling of concrete cracking, EUROMECH Colloquium 460 Numerical Modelling of Concrete Cracking, Innsbruck, Austria.
2. Armero F., Garikipati K. (1996). An analysis of strong discontinuities in multiplicative finite strain plasticity and their relation with the numerical simulation of strain localization in solids. Int. J. Solids & Structures, vol. 33, pp. 2863–2885.
3. Barragan B.E. (2002). Failure and toughness of steel fiber reinforced concrete under tension and shear, PhD. Thesis, UPC, Barcelona, Spain.

4. Belytschko T., Moës N., Usui S., Parimi C. (2001). Arbitrary discontinuities in finite elements. Int. J. for Num. Meth. Eng., vol. 50, pp. 993–1013.
5. Borja R.I. (2000). A finite element model for strain localization analysis of strongly discontinuous fields based ond standard Galerkin approximation. Comp. Meth. Appl. Mech. and Eng., vol. 190, pp. 1529–1549.
6. Feist C., Hofstetter G. (2006). An embedded strong discontinuity model for cracking of plain concrete, Comp. Meth. Appl. Mech. and Eng., vol. 195, pp. 7115–7138.
7. Gravouil A., Moës N., Belytschko T. (2002). Non-planar 3D crack growth by the extended finite element and level sets – Part II: Level set update. Int. J. for Num. Meth. in Eng., vol. 53, pp. 2569–2586.
8. Papoulia K.D., Sam C.H., Vavasis S.A. (2003). Time continuity in cohesive finite element modeling. Int. J. for Num. Meth. Eng., vol. 58, pp. 679–701.
9. Planas J., Elices M., Guinea G.V., Gómez F.J., Cendon D.A., Arbilla I. (2003). Generalizations and specializations of cohesive crack models. Engineering Fracture Mechanics, vol. 70, pp. 1759–1776.
10. Oliver J. (2000) On the discrete constituve models induced by strong discontinuity kinematics and continuum constitutive equations. Int. J. Solids & Structures, vol. 37, pp. 7207–7229.
11. Oliver J., Huespe A.E. (2004). Continuum approach to material failure in strong discontinuity settings. Comp. Meth. Appl. Mech. Eng., vol. 193, pp. 3195–3220.
12. Oliver J., Huespe A.E. (2004). Theoretical and computational issues in modelling material failure in strong discontinuity scenarios. Comp. Meth. Appl. Mech. Eng., vol. 193, pp. 2987–3014.
13. Oliver J., Huespe A.E., Samaniego E., Chaves EWV (2004). Continuum approach to the numerical simulation of material failure in concrete. Int. J. Num. Anal. Meth. Geomech., vol. 28, pp. 609–632.
14. Oliver J., Huespe A.E., Sánchez P.J.(2006). A comparative study on finite elements for capturing strong discontinuities: EFEM vs XFEM. Comput. Meth. Appl. Mech. Eng., vol. 195, pp. 4732–4752.
15. Oliver J., Huespe A.E., Blanco S., Linero D.L. (2006). Stability and robustness issues in numerical modeling of material failure in the strong discontinuity approach. Comput. Meth. Appl. Mech. Eng., vol. 195, pp. 7093–7114.
16. Ortiz M., Pandolfi A. (1999). Finite-deformation irreversible cohesive elements for three-dimensional crack-propagation analysis. Int. J. for Num. Meth. Eng., vol. 44, pp. 1267–1282.
17. Sancho J.M., Planas J., Cendón D.A., Reyes E., Gálvez J.C. (2007). An embedded crack model for finite element analysis of concrete – fracture. Eng. Fract. Mech., vol. 74, pp. 75–86.
18. Simo J.C., Oliver J., Armero F. (1993). An analysis of strong discontinuities induced by strain-softening in rate-independent inelastic solids. Comput. Mech., vol. 12, pp. 277–296.
19. Simo J.C., Oliver J. (1994). A new approach to the analysis and simulation of strain softening in solids. Proc. US-European Workshop on Fracture and Damage in Quasibrittle Structures, Z.P. Bazant, Z. Bittnar, M. Jirasek, M. Mazars (Eds.), Prague, Czech Republic.
20. Xu X.P., Needleman A. (1994). Numerical simulation of fast crack growth in brittle solids, J. Mech. Phys. Solids, vol. 42, pp. 1397–1434.

A 3D Cohesive Investigation on Branching for Brittle Materials

Rena C. Yu[1], Anna Pandolfi[2], and Michael Ortiz[3]

[1] E.T.S. de Ingenieros de Caminos, Canales y Puertos, Universidad de Castilla-La Mancha, 13071 Ciudad Real, Spain
[2] Dipartimento di Ingegneria Strutturale, Politecnico, 20133 Milano, Italy, pandolfi@stru.polimi.it
[3] Engineering and Applied Science Division, California Institute of Technology, Pasadena CA 91125

Summary. Recently, Fineberg and Sharon conducted dynamic crack propagation experiments in PMMA and soda lime glass [1, 2, 3, 4, 5, 6, 7]. They pointed out some notable features of micro-branching instabilities in brittle materials, and their experiments raised a considerable interest for the brittle fracture dynamics. In this paper we present some numerical results on brittle fracture obtained by using cohesive theories of fracture. In the numerical calculations, the branching instability is a natural outcome of the explicit formulation. The cohesive model captures the basic features of experiments, such as the transition of the crack surface from smooth to hackled with increasing energy flux, and the power-law functional form of the profile of the crack branches.

Key words: branching instabilities, cohesive models, brittle fracture, finite elements.

1 Introduction

During the last fifteen years, Fineberg and Sharon [1, 2, 3, 4, 5, 6, 7] conducted several experiments in brittle amorphous materials to study the propagation of cracks under dynamic loading. They observed that: (i) there exists a critical crack velocity, v_c, beyond which the dynamics of crack changes dramatically; (ii) for crack velocities $v < v_c$, a single main crack propagates straightforward, leaving a smooth fracture surface; (iii) for crack velocities $v > v_c$, small side branches sprout from the main crack, and a clear transition from a single crack to multiple cracks –branching instability– is observed. Usually, v_c is only a fraction of the Rayleigh wave speed v_R, the theoretical asymptotic crack speed for a single crack [8, 9]; for example, v_c is about $0.4v_R$ for PMMA.

Alain Combescure et al. (eds.), IUTAM Symposium on Discretization Methods for Evolving Discontinuities, 139–151.

Additionally, in their experiments Fineberg and Sharon showed that the critical velocity value v_c is independent of the specimen geometry and the loading rate; the branching trajectory is independent of the crack velocity and exhibits scaling behavior; and the branch length and the distance between branches are well modelled by a log-normal distribution.
The interest of the scientific community for dynamic fracture in brittle materials is testified by analytical works [10, 11] and numerical simulations with lattice models [17], molecular dynamics [12, 13] and finite element models [14, 15, 16].
We present here three dimensional finite element simulations of branching instability in brittle amorphous materials [18], obtained by combining cohesive theories of fracture [21] with the explicit simulation of fracture propagation [19, 20]. The cohesive theory of fracture is summarized in Sect. 2; the numerical simulations are described in Sect. 3, and finally some conclusions are drawn in Sect. 4.

2 Cohesive theory of fracture

Cohesive theories regard fracture as a progressive separation between two surfaces resisted by surface tractions. The relationship between the opening displacement $\boldsymbol{\delta} = [\![\mathbf{u}]\!]$ (strong displacement discontinuity between the two surfaces) and the resisting traction is defined by a cohesive law. In the simplest case, the cohesive law can be expressed as a function of two fracture parameters, for example the critical energy release rate G_c –the energy spent for a unit increase of the extension of the crack surface– and the tensile strength of the material σ_c:

$$\mathbf{t} = f(\boldsymbol{\delta}, G_c, \sigma_c) \tag{1}$$

By introducing a cohesive energy density per unit of undeformed surface, ϕ, function of $(\boldsymbol{\delta}, G_c, \sigma_c)$, the cohesive law is obtained by differentiating ϕ with respect to $\boldsymbol{\delta}$:

$$\mathbf{t} = \frac{\partial \phi(\boldsymbol{\delta}, G_c, \sigma_c)}{\partial \boldsymbol{\delta}} \tag{2}$$

Denoting with $\mathbf{n}$ the normal to the cohesive surface in the current configuration, we consider [22, 21] a simple class of mixed-mode cohesive laws accounting for tension-shear coupling, obtained by the introduction of an effective opening displacement:

$$\delta = \sqrt{\beta^2 \delta_S^2 + \delta_n^2} \tag{3}$$

where

$$\delta_n = \boldsymbol{\delta} \cdot \mathbf{n} \tag{4}$$

is the normal opening displacement and

$$\delta_S = |\boldsymbol{\delta}_S| = |\boldsymbol{\delta} - \delta_n \mathbf{n}| \tag{5}$$

is the magnitude of the sliding displacement. The parameter β assigns different weights to the sliding and normal opening displacements. The cohesive law (1) reduces to

$$\mathbf{t} = \frac{t}{\delta}(\beta^2 \boldsymbol{\delta}_S + \delta_n \mathbf{n}) \tag{6}$$

where t is a scalar effective traction. The expression of t follows from (3) and (6) as

$$t = \sqrt{\beta^{-2}|\mathbf{t}_S|^2 + t_n^2} \tag{7}$$

where $\mathbf{t}_S$ and t_n are the shear and the normal traction respectively. From this relation, we observe that β defines the ratio between the shear and the normal critical tractions and roughly defines the ratio of K_{IIc} to K_{Ic} of the material. The simple cohesive law adopted in the subsequent calculations is depicted in Fig. 1. Cohesive theories of fracture introduce into the material description

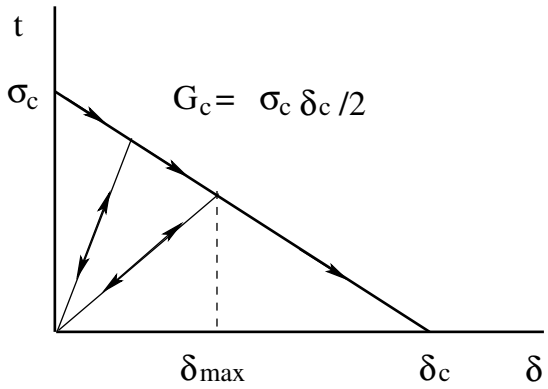


Fig. 1. Irreversible cohesive law adopted in the calculations, characterized by a linearly decreasing first loading envelop.

characteristic length and time scales [22, 23, 24]. The characteristic length ℓ_c is defined as:

$$\ell_c = \frac{EG_c}{\sigma_c^2} \tag{8}$$

where E is the elastic modulus of the material. The intrinsic length ℓ_c roughly estimates the size of the process zone, where the crack opening takes place. To avoid spurious mesh effects, the finite element size must resolve the process zone extension.

The characteristic time t_c is described as follows:

$$t_c = \frac{\rho c_L \delta_c}{2\sigma_c} \tag{9}$$

where ρ is the mass density, $c_L = \sqrt{E/\rho}$ is the longitudinal wave speed and δ_c the critical cohesive length. The intrinsic time t_c arises from the interplay

between material inertia and cohesive properties under dynamic conditions. Owing to the intrinsic time scale, cohesive models have been proved able to reproduce rate effects, although their definition requires only static material properties. The characteristic time provides a theoretical basis for the estimate of a critical strain rate for branching [18, 23, 24].
Cohesive approaches to evaluate branching instabilities have been previously presented by Xu and Needleman [15], and more recently by Falk, Needleman and Rice [16]. In their analysis, though, cohesive surfaces are initially embedded into the discretized continuum. In our simulations we use an automatic mesh adaption procedure, and explicitly simulated the nucleation and growth of cracks by inserting cohesive surfaces into the finite element model as requested by the numerical calculations [19, 20]. The adaptive insertion is driven by the attainment of a threshold value of the effective traction on the inter-element surfaces.

3 Simulation of dynamic branching

As aforementioned, we apply the cohesive theory of fracture accompanied with an insertion algorithm to model the branching instability in brittle materials. We adopted the experimental configuration by Fineberg and Sharon [1, 2, 3, 4, 5] and compare our numerical results with the experiments in PMMA specimens. We then repeat the same calculations for Homalite-100, another brittle material which shows dynamic branching instabilities. The mechanical properties for both materials are collected in Table 1.

Table 1. Mechanical parameters for PMMA and Homalite-100.

Material Property			PMMA	Homalite-100
Young Modulus	E	GPa	3.24	5.3
Poisson's ratio	ν		0.35	0.35
Material density	ρ	kg/m^3	1190	1230
Fracture energy	G_{Ic}	N/m	352.35	81
Cohesive stresses	σ_c	MPa	324	35
Characteristic length	L_c	μm	20	682
Characteristic time	t_c	μs	0.011	0.198

3.1 Experimental set-up and numerical treatment

The specimens used by Fineberg and Sharon [1, 3] were one edge pre-notched thin square plates, 50-100 mm wide, 200-400 mm long and 0.8-3.0 mm thick, made on cast PMMA. During the experiments, the crack tip position and

velocity were determined measuring the variation in the electrical resistance of a thin uniform aluminum layer evaporated onto a face of the specimen, see Fig. 2. The samples were quasi-statically loaded in mode-I applying a tensile

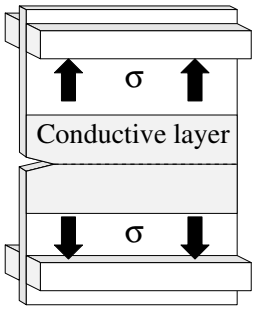


Fig. 2. Experimental setup of Fineberg and Sharon tests.

stress about half of the critical value to initiate the fracture. From that point the strain was increased by imposing a uniform displacement with a constant loading rate along the horizontal edges. Several loading rates were applied during the experimental program. It was observed that the transition from smooth to branching precedes the arrival at the crack tip of the transverse and longitudinal elastic waves generated at the onset of the fracture and reflected from the boundaries.

For simulations in PMMA, we need to accommodate the resolution of the

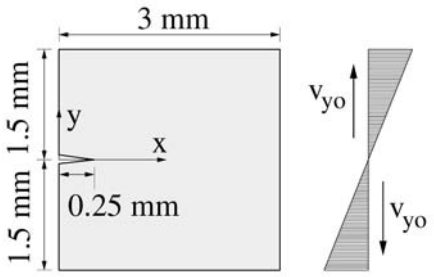


Fig. 3. Geometry of the simulated plate and distribution of the initial strain in the vertical direction.

process zone with the computational effort requested by a three dimensional model. The characteristic length of PMMA is particularly small, about 20 μm; therefore we maintain the geometry and loading conditions of the experimental program in a reduced scale. We modelled a thin plate $3 \times 3 \times 0.03$ mm with a 0.25 mm pre-notch, Fig. 3. To reproduce the experimental conditions, we apply an initial uniform strain rate γ (i.e. velocity $v_{yo}(y) = \gamma\, y$ in the vertical

direction. The dynamic loading is given by a constant velocity $v_y^{\pm} = \pm L\gamma/2$ (where $L = 1.5$ mm), imposed along the top and bottom edges of the specimen. We discretized the specimen with 10-node quadratic tetrahedral elements and

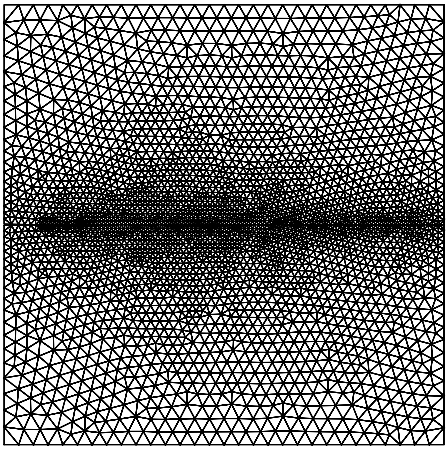


Fig. 4. Computational mesh adopted in the calculations, comprising 36,769 nodes and 18,162 10-node tetrahedra.

explicitly modelled the mid surface along the x-axis, where the main smooth crack develops under Mode I loading condition. In the zone surrounding the crack path the element size $h_{\min}$ has been taken equal to the characteristic length. The mesh progressively coarsens from the midsection to the edges. The initial computational mesh, see Fig. 4, consists of 36,769 nodes and 18,162 tetrahedrons.

As Sharon *et al.* observed [3], plasticity and rate effects do not affect the onset of branching instability, therefore we assume for the bulk an elastic behavior (Neo-Hookean finite elasticity model). In the calculations, the second-order accurate central difference algorithm is adopted to discretize the equation of motion in time.

3.2 The macro crack patterns

One of the most important findings in Fineberg and Sharon's experiments was that the crack speed and the crack structure change with the strain rate. In particular, at low strain rates, the crack speed remains below the critical velocity v_c and the crack surface appear prevalently smooth, Fig. 5(a). When the strain rate is increased and the mean crack speed reaches v_c, single short lateral branches develop and arrest immediately, Fig. 5(b). For higher loading rates, the mean crack speed remains beyond v_c and begins to oscillate in time; lateral branches proliferate and grow in longer paths, defining a precise side

crack profile, Figs. 5(c)-(d). In Fig. 6 we show the results of our numerical

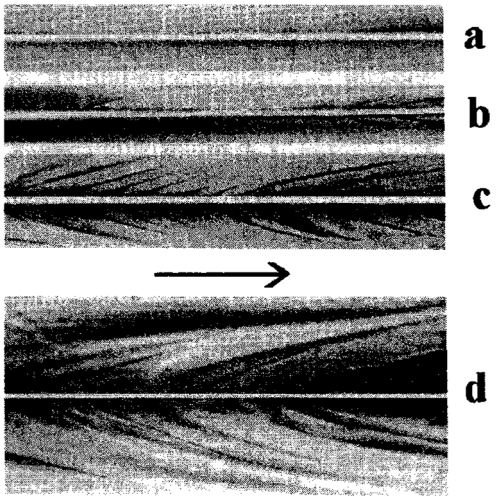


Fig. 5. Crack patterns observed by Fineberg and Sharon for different mean crack velocities (a) $v < 330$ m/s, (b) $v = 340$ m/s, (c) v =400 m/s and (d) v =480 m/s [1, 3].

simulations for three different strain rates: 500 s^{-1}, 1000 s^{-1} and 5000 s^{-1}. We were able to estimate the average velocity of the main crack by dividing the advance of the crack tip by the propagation time. The numerical crack velocities for the three loading rates were 214, 291 and 492 m/s respectively. Note that the third velocity was greater than the experimentally observed critical velocity for branching, i.e. 340 m/s. In Fig. 6, we plot the deformed configuration at the end of the analysis and the contour levels of the damage variable on the fracture surfaces. The damage D is a field variable, ranging from the values 0 and 1, defined *a posteriori* as the ratio between the consumed cohesive energy and the fracture energy G_c. For a low strain rate (500 s^{-1}) a single smooth and straight crack appears behind the wake, Fig. 6(a). When the strain rate increases up to 1000 s^{-1}, the crack develops in a smooth path only at the earlier stages of its growth, Fig. 6(b). Successively, the main crack attempts to branch out developing short lateral cracks. Most of the side branches only live up to one element and then arrest; only one side branch propagates parallel to the main crack for a short time interval (0.3 μs) and successively it merges again into the main crack. The main crack continues to grow as a single crack for a while. When a suitable amount of energy is stored, the crack branches out again. This sequence is repeated several times. For a higher strain rate (5000 s^{-1}), the single smooth crack develops for a shorter distance. In the initial stages of the crack propagation, small short-lived side branches form

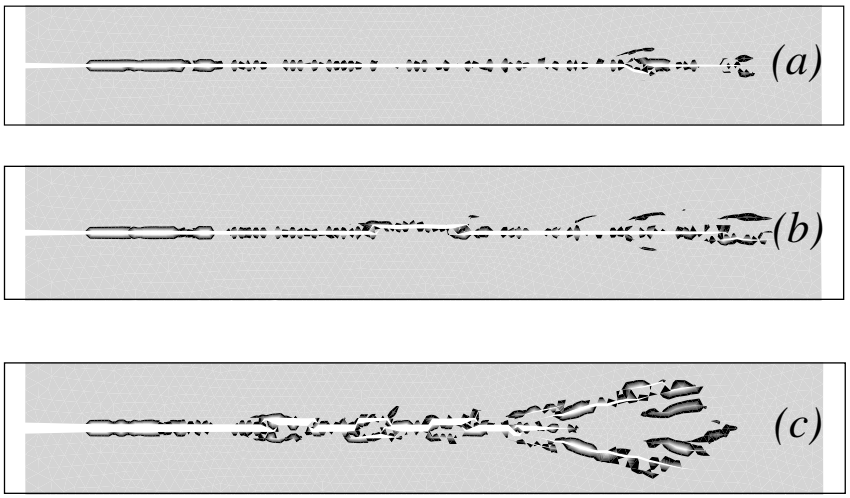


Fig. 6. Crack patterns as obtained through the numerical simulations for different loading rates at (a) 500 s^{-1} (b) 1000 s^{-1} and (c) 5000 s^{-1}. The figure shows the contour levels of the damage variable D. The average mean of the main crack velocity is estimated as (a) 214 m/s (b) 291 m/s and (c) 492 m/s.

and die, and the crack seems to be periodically hackled, Fig. 6(c). Finally, the main crack sprouts out into two long-lived branches, almost symmetric with respect to the main crack.

3.3 The micro branch profile

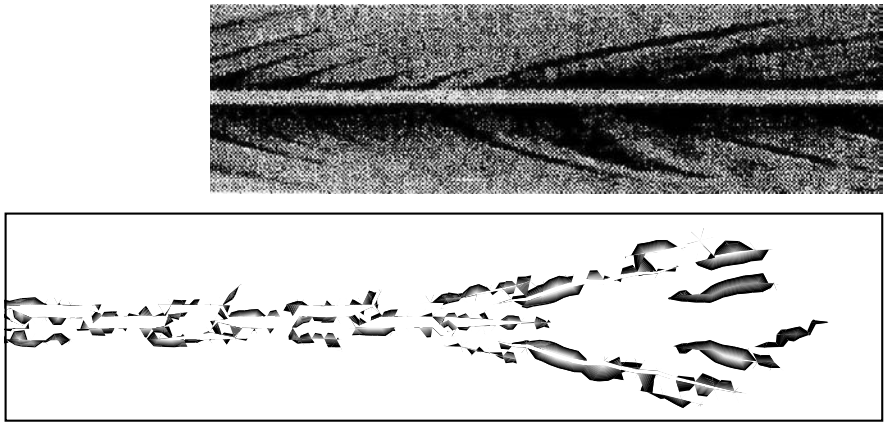


Fig. 7. Comparison between the experiments by Fineberg and Sharon [1, 3] (top) and detail of the present calculation (bottom) for high loading rate (5000 s^{-1}).

A close correspondence between experiments and simulations is also found in the branch pattern for higher loading rate, see Fig. 7. Note that in our numerical simulation the lateral branches develop from the main crack with a similar angle, and propagate following similar parabolic trajectories. Fineberg and Sharon [1, 2, 3, 4, 5, 6, 7] examined the shape and the structure of the micro branches, and described the mean profile $y(x)$ of the distributed branches with a power law equation:

$$y(x) = 0.2\, x^{0.7} \tag{10}$$

where x is the coordinate along the direction of the main crack. The experi-

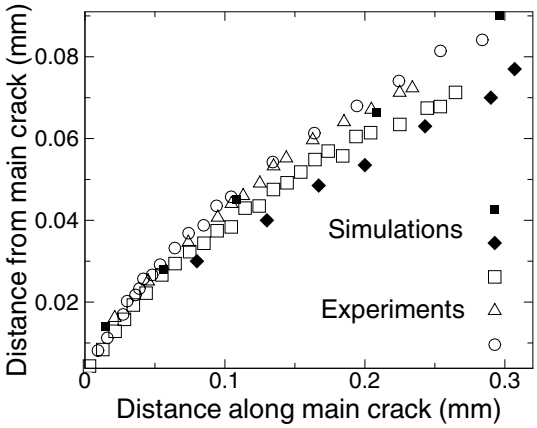


Fig. 8. Crack profiles: comparison between experiments [1, 3] and numerical simulations.

ments proved that the initial slope of the branch trajectory is independent of the crack velocity; anyway the crack profiles diverge in later times, accordingly to the crack speed. In particular, the power law relation holds for velocities within 10 percent of v_c, but the coefficient 0.2 requires a slight adjustment for higher velocities. In Fig. 8 we compare the branch profile obtained in the simulations (for strain rates equal to 1000 s^{-1} and 5000 s^{-1}) with the experiments. The best fitting power interpolation for the two sets of numerical data matches the exponent 0.7 of the profile power law; the coefficient ranges between 0.17 and 0.2, testifying to a very good agreement between our simulation and experiments.

3.4 The energy dissipation mechanism

Sharon *et al.* [3, 4] pointed out that beyond the onset of instability, any amount of energy supplied to the crack that exceeds the amount required

for the smooth crack propagation is not used to increase the crack tip speed, but in creating new fracture surfaces. As a consequence, the total fracture area of micro-branches is nearly an order of magnitude larger than the one of a single smooth crack; thus the formation of such surfaces absorbs a much larger amount of energy. This explains why the theoretical limiting velocity v_c cannot be experimentally realized in brittle materials: a crack prefers to dissipate energy by creating new surfaces via multiple parallel cracks rather than by accelerating. In order to have a better evaluation of the energy dissi-

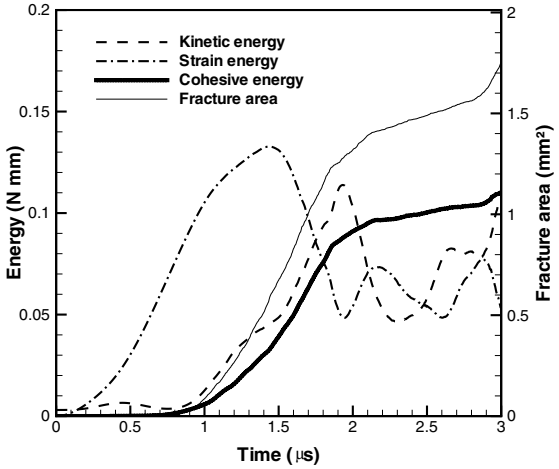


Fig. 9. Top: Energy dissipation and fracture area history. Bottom: Cohesive energy and kinetic energy rate plotted against the strain/deformation energy history.

pation mechanism, in Fig. 9, we plot the energy history and the fracture area history obtained numerically for PMMA loaded at 5000/s strain rate. The strain energy is computed by integrating the strain energy density of each finite element of the mesh. The kinetic energy is computed by summing the contribution of the kinetic energy of all the nodes of the mesh. The fracture energy is computed by integrating the cohesive energy density on all the cohesive surfaces surfaces. Finally, the sum of the area of the completely fractured cracks defines the fracture area. The strain energy globally increases in the initial stages of the loading process. As more energy is supplied the system, new surfaces are created: the cohesive energy and the total fracture area start to increase. A steep increase of the fracture area and the cohesive energy denotes the onset of branching. At this point, the deformation energy starts to decrease.

Fig. 10 combines the strain energy history with the cohesive energy rate history and kinetic energy rate history. The rate plots are obtained by numerical

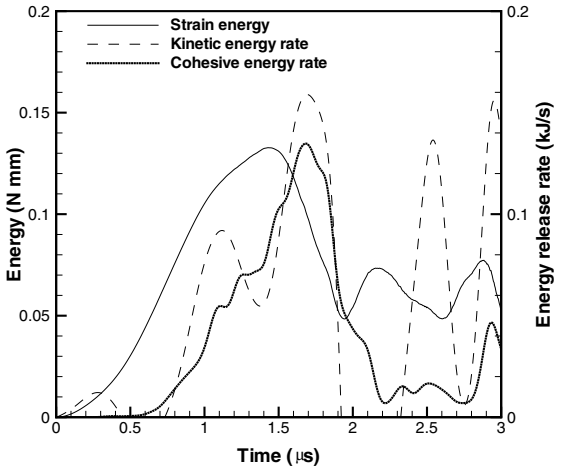


Fig. 10. Top: Energy dissipation and fracture area history. Bottom: Cohesive energy and kinetic energy rate plotted against the strain/deformation energy history.

differentiation of the plots in Fig. 9. Both the kinetic energy and the cohesive energy rate curves show a peak shortly after the peak in the deformation energy. This indicates that, through the branching mechanism, the accumulated deformation energy is partially dissipated into cohesive energy and partially transformed into kinetic energy. This observation is in keeping with the hypothesis that the formation of side branches does dissipate more energy than the propagation of a single smooth crack.

3.5 Another example of branching instability

In order to verify the ability of the method to capture branching instability in different brittle materials, we repeated the same numerical calculation (using the same mesh) for Homalite-100, whose characteristic length is 682 μm. We observed that the numerical results for a loading rate of 1000 s^{-1} were similar to the ones obtained for the PMMA case at 5000 s^{-1}, although the length of the hackled region was smaller and the long-lived branches appeared earlier, Fig. 11. Due to the wave reflection, spurious cracks develop close to the loaded boundaries and far from main trajectory of the crack; a straight crack also starts in the mid-plane form the side opposite to the original pre-notch. One of the two branches tries to coalesce with the main crack during the final stages of the process.

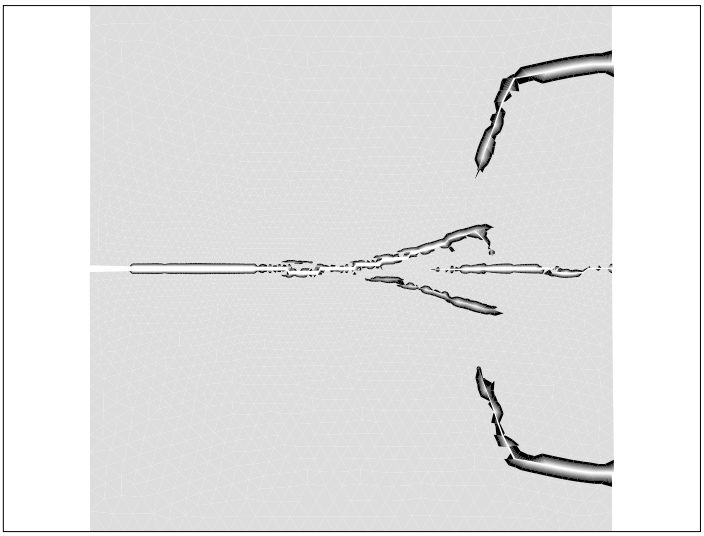


Fig. 11. Crack pattern obtained with the numerical simulation in Homalite-100 plate.

4 Conclusions

We presented numerical simulations of branching instability in brittle materials. We used a three dimensional finite element model and cohesive theories of fracture. The results for PMMA specimens are compared with the main features of branching observed in the experiments by Fineberg and Sharon. The branching instability is a natural outcome of the explicit formulation of crack propagation. Our model captures the transition of the crack surface from smooth to hackled with increasing energy flux and the power-law functional form of the crack branch profile. The analysis of the energy history confirms that brittle materials prefer to dissipate energy through the formation of side branches rather than increasing the propagation speed of a single main crack.

Acknowledgments

AP acknowledges the support of the Italian MIUR PRIN-2005 programme, Protocol no. 2005085973. RCY thanks the Spanish *Ministerio de Educación y Ciencia* for the funding through the *Programa Ramón y Cajal*.

References

1. Fineberg J, Gross SP, Marder M, Swinney HL (1992) Physical Review Letters 67:457–460

2. Fineberg J, Gross SP, Marder M, Swinney HL (1992) Physical Review B 45:5146–5154
3. Sharon E, Gross AP, Fineberg J (1995) Physical Review Letters 74:5096–5099
4. Sharon E, Gross AP, Fineberg J (1996) Physical Review Letters 76:2117–2120
5. Sharon E, Fineberg J (1996) Physical Review B 54:7128–7139
6. Sharon E, Fineberg J (1998) Philosophical Magazine B 78:243–251
7. Sharon E, Fineberg J (1999) Letters to Nature, 397:333–335
8. Freund LB (1990) Dynamic Fracture Mechanics. Cambridge Univ. Press, New York.
9. Ravi-Chandar K, Knauss WG (1984) International Journal of Fracture, 26:65–80
10. Marder M, Liu X (1993) Physical Review Letters 71:2417–2420
11. Marder M, Gross S (1995) Journal of the Mechanics and Physics of Solids 43:1–48
12. Abraham FF, Brodbeck D, Rafey RA, Rudge WE (1993) Physical Review Letters 73:272–275
13. Zhou SJ, Lomdahl PS, Thomson R, and Holian BL (1996)Physical Review Letters 76:2318–2321
14. Johnson E (1992) International Journal of Fracture 55:47–63
15. Xu XP, Needleman A (1994) Journal of the Mechanics and Physics of Solids 42:1397–1434
16. Falk M, Needleman A, Rice JR (2001) Journal de Physique IV 5.43–5.50.
17. Martin T, Espanol P, Rubio MA, Zuniga I (2000) Physical Review E 61:6120–6131
18. Yu RC (2001) Three-dimensional Cohesive Modelling of Impact Damage of Composites. PhD thesis, California Institute of Technology, Pasadena, CA
19. Pandolfi A, Ortiz M (1999) Engineering with Computers 14:287–308
20. Pandolfi A, Ortiz M (2002) Engineering with Computers 18:148–159
21. Ortiz M, Pandolfi A (1999) International Journal for Numerical Methods in Engineering 44:1277–1282
22. Camacho GT, Ortiz M (1996) International Journal of Solids and Structures 33:2899–2938
23. Ruiz G, Ortiz M, Pandolfi A (2000) International Journal for Numerical Methods in Engineering 48:963–994
24. Ruiz G, Pandolfi A, Ortiz M (2001) International Journal for Numerical Methods in Engineering 52:97–120

PARTITION-OF-UNITY BASED FINITE ELEMENT METHODS

On Applications of XFEM to Dynamic Fracture and Dislocations

Ted Belytschko[1], Jeong-Hoon Song[2], Hongwu Wang[3], and Robert Gracie[4]

[1] Theoretical and Applied Mechanics
Northwestern University
2145 North Sheridan Road Evanston, IL 60208-3111
tedbelytschko@northwestern.edu
[2] j-song2@northwestern.edu
[3] h-wang7@northwestern.edu
[4] rgracie@northwestern.edu

Summary. The application of the extended finite element method, which can model arbitrary discontinuities in finite elements to dynamic fracture and dislocations, is described. The dynamic fracture methodology is applied to a problem of crack branching and compared to element deletion and interelement crack methods. The dislocation method directly models the dislocation as a tangential discontinuity. This makes the method readily applicable to problems with interfaces and anisotropy.

Key words: extended finite element method; dynamic fracture; dislocations

We describe some recent studies in the extended finite element method (XFEM) and related methods for the treatment of dynamic crack propagation and a new approach for treating dislocations by XFEM. The extended finite element method is a methodology for modeling cracks of arbitrary geometry in a finite element method without remeshing. It originated in Belytschko and Black [1] and Moës et al. [2], and was combined with level sets in Stolarska et al. [3] and Belytschko et al. [4]. The method can be viewed as a partition of unity method [5], but in fact, in the treatment of discontinuities by this method, a partition of unity is never constructed. Instead, a discontinuous function that only spans one element and vanishes at the edges is constructed. In the new approach to dislocations, this contrast with a partition of unity is even more apparent, for the enrichment is a single function which introduces a discontinuity along the glide line.

Dynamic crack propagation is an application domain for which XFEM is particularly suitable because the most prevalent method for treating crack growth, remeshing, is very awkward for these problems. In most dynamic crack propagation problems, the crack advances over a large part of the mesh,

Alain Combescure et al. (eds.), IUTAM Symposium on Discretization Methods for Evolving Discontinuities, 155–170.

so that remeshing would need to be performed many times. Remeshing is especially daunting for fragmentation simulations, since such simulations would entail very extensive remeshing. Furthermore, the continuity of the solution is quite important in any numerical model, and even with excellent projection schemes, significant discontinuities can be introduced in the velocities, stresses and displacements by remeshing.

XFEM was first applied to dynamic crack propagation in Belytschko et al. [6]. Subsequently, Song et al. [7] used the Hansbo and Hansbo [8] approach which uses the same basis functions; see Areias and Belytschko [9]. Réthoré et al. [10, 11] improved the stability of the explicit time integration scheme and developed an energy conserving dynamic crack propagation algorithm with XFEM. Remmers et al. [12] have introduced an interesting variant of XFEM where they initiate a crack by injecting discontinuities that span three elements at a time.

Here, we will compare three methods for dynamic crack propagation:

1. the extended finite element method
2. the interelement crack method
3. the element deletion method

Dynamic crack propagation is a difficult area in which to benchmark different methods because there are no analytical solutions which can readily be compared to numerical results. Therefore we will make comparisons to experimental results and examine how well the three methods reproduce various aspects of a selected experiment.

Section 3 describes a new method for modelling dislocations that was recently reported by Gracie et al. [13]. In contrast to most previous methods, in the XFEM approach to dislocations, superposition is not used. Instead, an interior dislocation is introduced. This is closely related to the Volterra concept of a dislocation, in which a glide dislocation is modelled by cutting a solid along the glide plane, sliding it by the Burgers vector, reattaching the solid along the glide plane and then determining the stresses in the solid. In our approach, the process is modelled by introducing an interior discontinuity whose magnitude is given by the Burgers vector and then computing the displacements in the rest of the body.

This XFEM method for dislocations can readily be implemented in standard finite element programs, since it only involves the computation of external forces that result from the interior discontinuities. More important, the method is applicable to many highly relevant problems that could not be tackled by superposition based methods: problems with interfaces, grain boundaries, and material anisotropy.

1 Dynamic Crack Propagation

We first briefly describe the extended finite element method, the interelement crack method, and the element deletion method, which is also often called element killing and element erosion methods.

The implementation of XFEM used in this study is based on Song et al. [7], which differs in performance from Belytschko et al. [6] only in its omission of partially cracked elements. The key idea in the formulation of this method is that the displacement field incorporates the discontinuity as additional terms in the displacement approximation.

Let the crack surface be defined by $f(\mathbf{X}) = 0$ and $g(\mathbf{X}, t) > 0$ where $\mathbf{X}$ are the reference coordinates; these implicit function $f(\mathbf{X})$ and $g(\mathbf{X}, t)$ are often called level sets. Then the FE displacement field approximation is

$$\mathbf{u}^h(\mathbf{X}, t) = \sum_I N_I(\mathbf{X})\{\mathbf{u}_I(t) + H(f(\mathbf{X}))H(g(\mathbf{X}, t))\mathbf{q}_I(t)\} \tag{1}$$

where $N_I(\mathbf{X})$ are the standard finite element shape functions, $\mathbf{u}_I(t)$ are the nodal displacements and $\mathbf{q}_I(t)$ are additional nodal parameters that govern the strength of the discontinuity.

In XFEM, the enrichment is injected when a criterion for crack nucleation or crack growth is met. In the solutions reported here, we used loss of material stability criterion (i.e. loss of hyperbolicity), as in [6, 7]. This conditional injection is one of the major drawbacks of the method which is discussed in more details later. The derivation of the equations of motion for the finite element model then follows the usual procedure and yields

$$\mathbf{M}_I\ddot{\mathbf{u}}_I + \mathbf{f}_I^{int} = \mathbf{f}_I^{ext} \tag{2}$$

See Belytschko et al. [6] and Song et al. [7] for details.

In the interelement crack method, the crack is modeled by separating along the element edges. This involves adding extra nodes, see Figure 1. Two approaches are used. In the original formulation of Xu and Needleman [14], all elements are separated from the beginning of the simulation. The edges are mechanically joined by cohesive laws of the form

$$\begin{aligned} T_n &= -\frac{\phi_n}{\Delta_n} e^{(-\delta_n/\Delta_n)}\{\frac{\delta_n}{\Delta_n} e^{(-\delta_n^2/\Delta_n^2)} + \frac{1-q}{r-1}[1 - e^{(-\delta_n^2/\Delta_n^2)}](r - \frac{\delta_n}{\Delta_n})\} \\ T_t &= -\frac{\phi_n}{\Delta_n}(2\frac{\Delta_n}{\Delta_t})\frac{\delta_t}{\Delta_t}\{q + \frac{r-q}{r-1}\frac{\delta_n}{\Delta_n}\} e^{(-\delta_n/\Delta_n)} e^{(-\delta_t^2/\Delta_t^2)} \end{aligned} \tag{3}$$

where $\mathbf{T}$ is the traction across the interelement surface, n and t denote the normal and tangential components respectively, $\boldsymbol{\delta}$ is the displacement jump across the cohesive surface interface, ϕ is the cohesive potential function, and Δ is a characteristic length; for details, see [14]. In the Ortiz and Camacho approach [15], elements are allowed to separate along edges only when a criterion is met or the element edges are contiguous to a crack tip. The cohesive law used in Camacho and Ortiz is

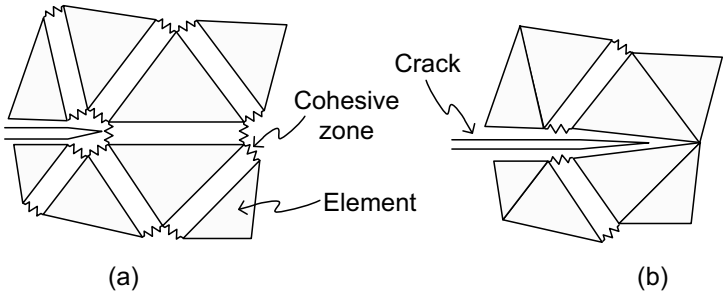


Fig. 1. Schematic showing of the Xu-Needleman method [14] and (b) the Camacho and Ortiz method [15].

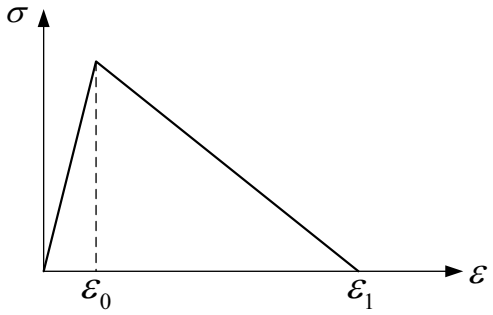


Fig. 2. Schematic showing of a stress-strain curve.

$$T_n = e\frac{\sigma_c}{\delta_c}\beta^2\delta_s e^{-\delta/\delta_c}$$
$$T_t = e\frac{\sigma_c}{\delta_c}\delta_n e^{-\delta/\delta_c} \tag{4}$$

where σ_c and δ_c are the maximum cohesive traction and critical opening displacement, respectively; for details, refer [15]. Similar laws are used in Ortiz and Pandolfi [16]. Once separation occurs, a cohesive law such as Eq. 3 or Eq. 4 is applied between the opposing edges, and the crack automatically chooses its path by solving the momentum equation that takes into account the traction forces on the cohesive edges.

The underlying concept in the so-called element deletion method is to employ a stress-strain law in which the stress goes to zero for sufficiently large strains, an example of such a stress-strain law for one dimension is shown in Figure 2. The dissipation in an element that is associated with this algorithm is then equated to the surface energy of a crack passing through the element by modifying the stress strain law. This energy consistency renders solutions relatively mesh size independent. Generally, no information about

the orientation of the crack surface is included, so it is best to use square or nearly square elements.

To illustrate the energy equivalence in two dimensional problems for the stress-strain law shown in Figure 2, the upper strain limit ϵ_1 is scaled so that

$$G_f h^e = \frac{1}{2} E \epsilon_0 \epsilon_1 A^e \tag{5}$$

where G_f is the fracture energy, h^e is a characteristic dimension (the length of a side for a square element), and A^e is the area of the element, unit thickness is assumed. While the method is called an element deletion method, the element is not deleted, but instead the stress in the element is set to zero. Some programs, such as LS-DYNA [17], remove the mass of the element from the global mass matrix to eliminate the inertia effects for these elements, but this seems unwarranted, since the mass does not disappear. We have retained the mass, but for the problems considered here, it makes little difference.

In the calculations reported here, we used the continuum damage model proposed by Lemaitre [18]. The damage evolution law is given by

$$D(\bar{\epsilon}) = 1 - (1 - A)\epsilon_0 \bar{\epsilon}^{-1} - A e^{-B\langle \bar{\epsilon} - \epsilon_0 \rangle} \tag{6}$$

where D is the damage parameter which can have values 0 to 1, $\bar{\epsilon}$ is the effective strain, A and B are material parameters, ϵ_0 is the strain threshold, and $\langle \cdot \rangle$ is the MacCaulay bracket. The constitutive relation is given by:

$$\sigma_{ij} = (1 - D) C_{ijkl} \epsilon_{kl} \tag{7}$$

where C_{ijkl} is the elastic modulus of the undamaged material, and σ_{ij} and ϵ_{kl} are stress and strain components, respectively.

All calculations were performed with 3 node triangular elements or 4 node quadrilateral elements in which one point integration rule is used. Both structured and unstructured meshes were used as noted; all calculations were made with explicit time integration. The time step was usually a very small fraction of the stable time step. This appears to be unavoidable in crack propagation problems at this time.

2 Benchmark Problem

We have chosen as a benchmark problem crack growth in a prenotched glass sheet. The setup of the problem is shown in Figure 3. As can be seen, a tensile stress $\sigma_y = 1$ MPa is applied at the top and bottom surfaces; the time history of the load is a step function. Experiment on specimens of similar dimensions have been reported by [19–23].

In these experiments, a crack starts growing at the notch and propagates to the right, generally with increasing speed. At a certain point, the crack branches into at least two cracks (some experiments show more branches). It

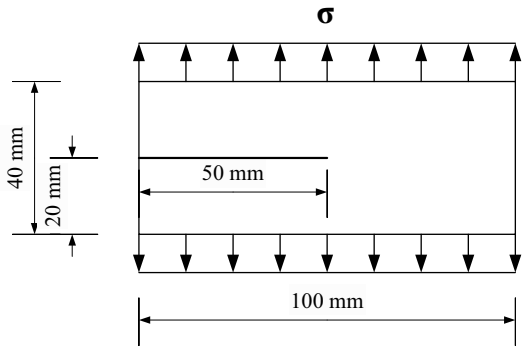


Fig. 3. Setup of a dynamic crack branching problem.

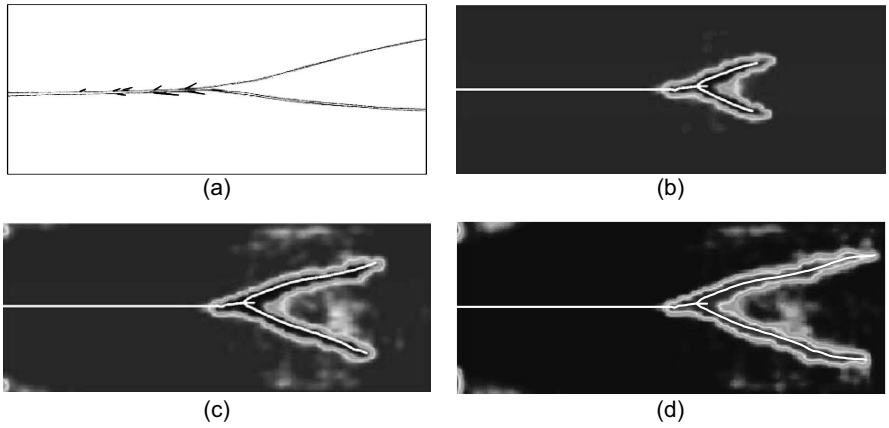


Fig. 4. Comparison of the XFEM result with the experiment:(a) a sketch of the experiment crack paths reported by Ramulu and Kobayashi [19]; XFEM plots of crack branching and damage evolution at (b) t=39.29 μs (c) t=46.14 μs and (d) t=55.93 μs.

is also noteworthy that as shown in Figure 4(a), prior to the major branching, minor branches appear to start from the major crack but only grow to small lengths.

Figures 4(b)-(d) show the evolution of the crack for a structured mesh at three different time steps as computed by XFEM. It can be seen that the major features of the crack path observed in the experiment are reproduced quite well. One phenomenon observed in the experiments that is missing in the XFEM calculations are the tentative branches before the major branch emerges, but the general shape of the crack branching paths agree. The results for an unstructured mesh are almost identical.

Fig. 5. Final crack path of the element deletion method with a 200×81 quadrilateral mesh at different time steps: (a) $t = 30.56\ \mu s$ and (b) $t = 65.62\ \mu s$

Figure 5 shows the results for the element deletion method at two different times. It can be seen that the crack continues in a straight path without branching. The calculation shown was made with a 200×81 mesh, but we have also used 50×21 and 100×41 meshes and obtained a straight crack with these meshes. One would expect the finite element solution to exhibit an increase in the tensile strain adjacent to the crack as the velocity increases, which would trigger crack branching. In the simulation by the element deletion method, the stress adjacent to the crack tip does increase, but it never becomes large enough to initiate crack propagation in the lateral direction.

Figures 6(a)-(b) and (c)-(d) show the final paths obtained by the interelement crack method [15] for 76×30 and 152×60 structured cross-triangular meshes, respectively. In these calculations, $\sigma_c = 2.8$ MPa and $\delta_c = 3.9 \times 10^{-7}$ m were used for Eq. 4. The interelement method predicts crack branching but the branching point depends somewhat on mesh refinement.

We should make some remarks about the criteria in the interlement method and in XFEM. In the interelement method, the crack propagation was modeled strictly by separation of the element edges. The only aspects of the algorithm that limits the evolution of the crack is the transition from edges that share nodes (and thus can not separate) to edges that have duplicated nodes (and thus can separate). In XFEM, the crack growth was governed by the material stability criterion, which is a property of the constitutive equation used (in these computations, Eqs. 6 and 7). Crack branching was initiated when the two polarization angles from the material stability analysis differed from that of the previous crack direction. When this occurred, the crack was also injected into the adjacent elements. Unless the latter step was taken, crack branching could not be modeled accurately with XFEM with low order elements.

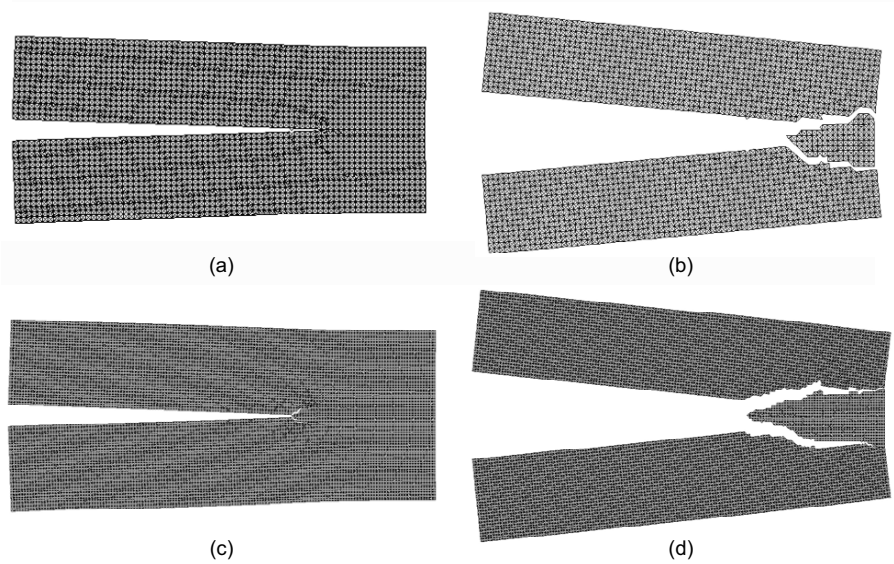


Fig. 6. Crack path of the Camacho and Ortiz's method [15] with two structured cross-triangular meshes: the crack path of a 76×30 structured cross-triangular mesh at (a) $t = 48.38\ \mu s$; (b) $t = 100.00\ \mu s$ and the crack path of a 152×60 structured cross-triangular mesh at (c) $t = 48.09\ \mu s$; (d) $t = 100.0\ \mu s$. The deformed shapes are magnified by 90 times.

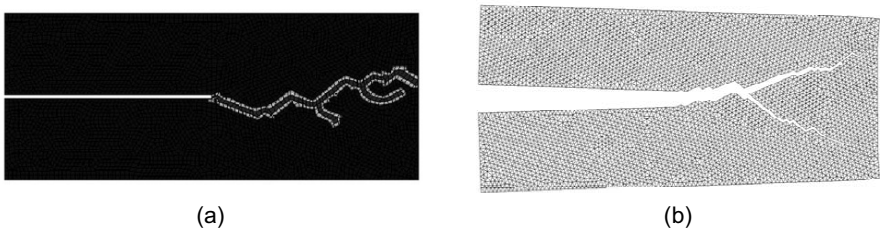


Fig. 7. Final crack path for unstructured mesh($h^e_{avg} = 1.0$ mm): (a) the result of the element deletion and (b) the results of the Camacho and Ortiz's method [15]. The deformed shapes are magnified by 90 times.

3 Dislocation Modeling by XFEM

We consider an edge dislocation, as illustrated in Figure 8. The geometry is described by an affine function of the coordinates, where

$$f^{\alpha}(\mathbf{x}) = \alpha_0 + \alpha_j x_j = 0 \tag{8}$$

and repeated indices denote summation. The active slip plane is described by $g^{\alpha}(\mathbf{x}) > 0$.

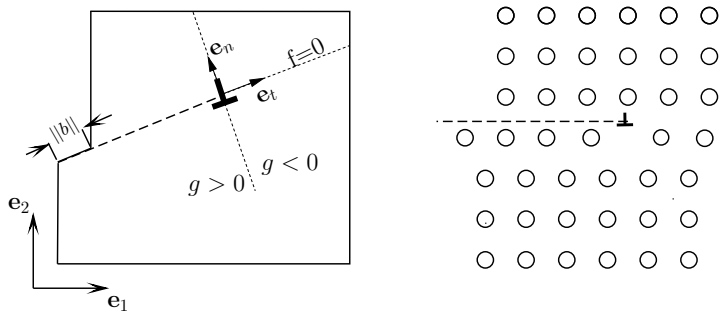


Fig. 8. Description of an edge dislocation by functions $f(\mathbf{x})$ and $g(\mathbf{x})$. Dashed line represents the glide plane, $\mathbf{b}$ is Burgers vector.

The displacement approximation with tangential enrichment for an edge dislocation with Burgers vector $b^\alpha \mathbf{e}_t^\alpha$ has the form

$$\mathbf{u}(\mathbf{x}) = \sum_{I\in\mathcal{S}} N_I(\mathbf{x})\,\mathbf{u}_I + \sum_{\alpha=1}^{n_D} b^\alpha \mathbf{e}_t^\alpha \sum_{J\in\mathcal{S}^\alpha} N_J(\mathbf{x})\left(H\left(f^\alpha(\mathbf{x})\right) - H_J\right)\bar{H}\left(g^\alpha(\mathbf{x})\right) \tag{9}$$

where $H_J = H\left(f^\alpha(\mathbf{x_J})\right)$, $\mathcal{S}$ is the set of all nodes, $\mathcal{S}^\alpha$ is the set of enriched nodes to be defined later, n_D is the number of dislocations, N_I are the standard finite element shape functions, $\mathbf{u}_I$ are the nodal displacements, ${\mathbf{e}^\alpha}_t$ is a unit tangent along the glide plane α, $\bar{H}(z) = H(z) + 1/2$ and $H(z)$ is the symmetrized Heaviside function given by

$$H(z) = \begin{cases} -1/2, & z < 0 \\ 1/2, & z > 0 \end{cases} \tag{10}$$

We call the second term in Eq. 9 an enrichment. The nodes that are enriched by the tangential step function, i.e. those in the set $\mathcal{S}^\alpha$, are shown in Figure 9. It can be seen that $\mathcal{S}^\alpha$ consists of all nodes of elements with at least two edges cut by glide plane α.

From the principle of virtual work or minimum potential energy, the following discrete equations are obtained

$$\mathbf{K}_{uu}\mathbf{u} + \mathbf{K}_{ub}\boldsymbol{\beta} = \mathbf{f}^{ext} \tag{11}$$

where $\mathbf{u} = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n]^T$ and the vector $\boldsymbol{\beta} = [\mathbf{b}^1, \mathbf{b}^2, \ldots, \mathbf{b}^{n_d}]^T$ consists of the slips along the glide planes. The stiffness matrix partitions are given by

$$\mathbf{K}_{IJ}^{uu} = \int_{\Omega/\Gamma_d} {\mathbf{B}^{\mathbf{T}}}_I \mathbf{C} \mathbf{B}_J d\Omega \tag{12}$$

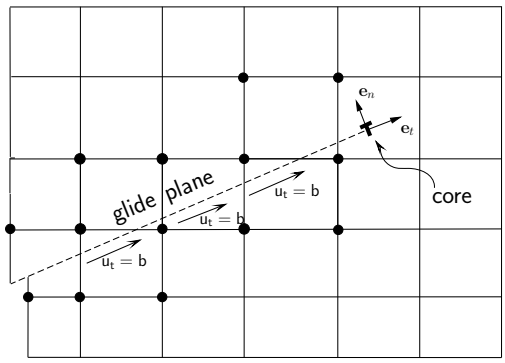


Fig. 9. Illustration of the tangential enrichment scheme. Dashed line represents the glide plane and black circles represent nodes of the set $\mathcal{S}^\alpha$ that are enriched.

$$\mathbf{K}_{I\alpha}^{ub} = \int_{\Omega/\Gamma_d} \mathbf{B}^{\mathbf{T}}{}_I \mathbf{C} \mathbf{D}_\alpha d\Omega \tag{13}$$

where

$$\mathbf{B}_I = \begin{bmatrix} N_I(\mathbf{x})_{,x} & 0 \\ 0 & N_I(\mathbf{x})_{,y} \\ N_I(\mathbf{x})_{,y} & N_I(\mathbf{x})_{,x} \end{bmatrix} \tag{14}$$

and

$$\mathbf{D}_\alpha = \sum_{I \in \mathcal{S}^\alpha} \begin{bmatrix} \left(H\left(f^\alpha(\mathbf{x})\right) - H_I^\alpha\right) N_{I,x} \left(\mathbf{e}_t^\alpha \cdot \mathbf{e}_x\right) \\ \left(H\left(f^\alpha(\mathbf{x})\right) - H_I^\alpha\right) N_{I,y} \left(\mathbf{e}_t^\alpha \cdot \mathbf{e}_y\right) \\ \left(H\left(f^\alpha(\mathbf{x})\right) - H_I^\alpha\right) N_{I,y} \left(\mathbf{e}_t^\alpha \cdot \mathbf{e}_x\right) + \left(H\left(f^\alpha(\mathbf{x})\right) - H_I^\alpha\right) N_{I,x} \left(\mathbf{e}_t^\alpha \cdot \mathbf{e}_y\right) \end{bmatrix} \tag{15}$$

where $H_I^\alpha = H\left(f^\alpha(\mathbf{x_I})\right)$. The nodal displacements are given by

$$\mathbf{u} = \mathbf{K}_{uu}^{-1} \left(\mathbf{f}^{ext} - \mathbf{K}_{ub}\boldsymbol{\beta}\right) \tag{16}$$

Since $\mathbf{K}_{uu}$ is independent of the location, number and geometry of the dislocations, it does not change for a given mesh as the dislocations move or as new dislocations are nucleated. In a dislocation dynamics simulation, $\mathbf{K}_{uu}$ needs only to be inverted once for the entire simulation. The Burgers vectors are given in dislocation dynamics simulations at any given time, so $\boldsymbol{\beta}$ is known and the effect of the dislocations is entirely on the right hand side of Eq. 16. Therefore standard commercial software can easily be adapted to implement this approach.

To illustrate the capabilities of the method we consider a 10^{-4} cm $\times$ 10^{-4} cm domain containing a dislocation core with a horizontal glide plane, see Figure 10. Along the boundary ABCD we apply displacement boundary conditions corresponding to the exact solution for the infinite domain. The

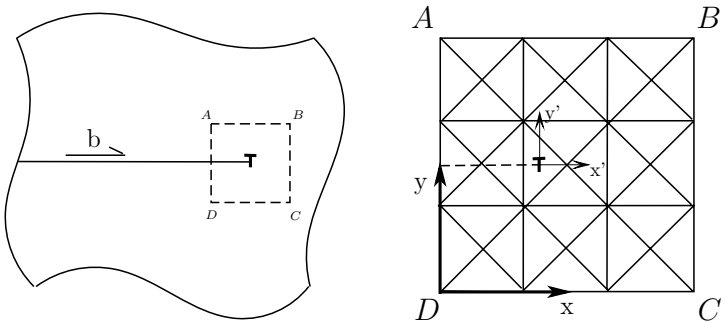


Fig. 10. Illustration of the sub-domain ABCD of an infinite body used to simulate a dislocations in an infinite domain.

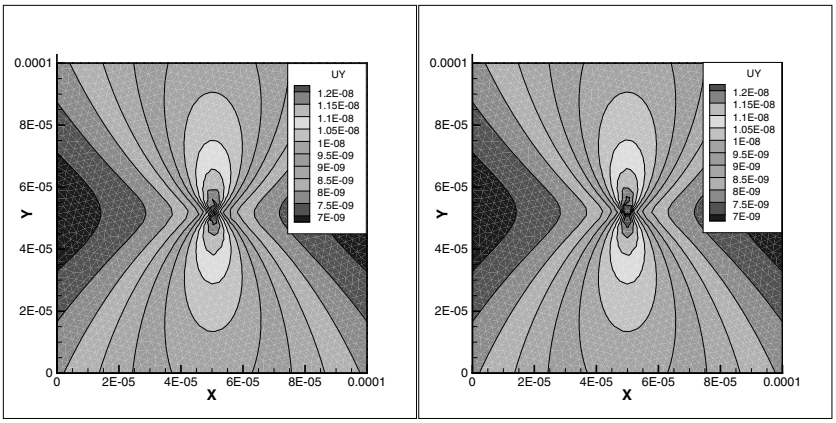


Fig. 11. Displacement in the y-direction in cm. On the right is the exact field; on the left are the results from XFEM.

elastic modulus, Poisson's ratio and the magnitude of the Burgers vector are 121.41 GPa, 0.34, and 8.551×10^{-8} cm, respectively. An unstructured triangular mesh with about 3600 elements is used. This corresponds to an element edge length of about $100\,\|\mathbf{b}\|$. The numerical solution is compared to the analytical solution for an edge dislocation in an infinite domain, given by Hirth and Lothe [24], in Figure 11.

Next we consider an edge dislocation in a semi-infinite domain near a free surface, as shown in Figure 12. The free surface is located at $x = 0$ and the domain is assumed to be semi-infinite and to occupy the domain $x > 0$. The dislocation is located a distance of $L = 0.5 \times 10^{-4}$ cm from the free surface and it is assumed that the glide plane is perpendicular to the free surface, along $y = 0$. The analytical solution to this problem is given by

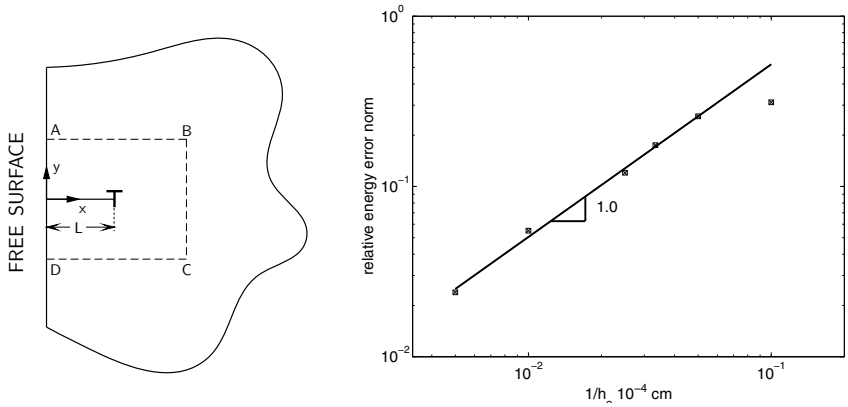


Fig. 12. Left, edge dislocation in a semi-infinite domain, near a free surface. Subdomain ABCD is the numerical simulation domain. Right, convergence of the relative energy error norm with decreasing element size, h_e. The convergence rate is 1.0

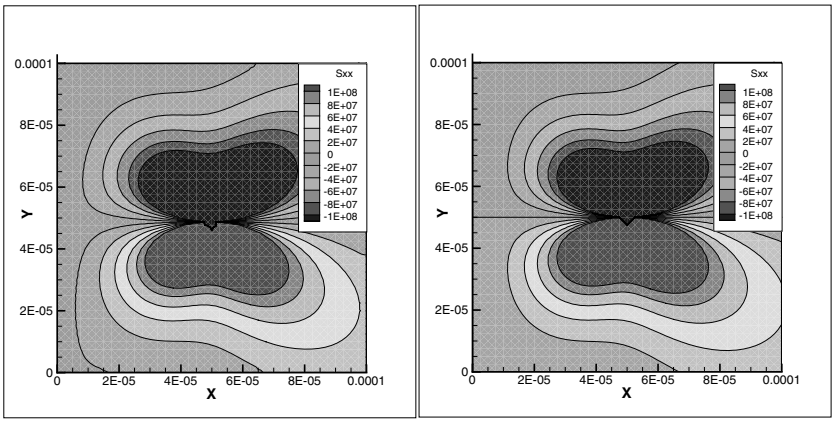


Fig. 13. Stress σ_{xx} in dynes/cm^2 for a dislocation in a semi-infinite domain. On the right is the exact field; on the left are the results from XFEM.

Head, [25]. We solve the problem on a subdomain ABCD, as in Figure 12, with dimensions 10^{-4} cm $\times$ 10^{-4} cm. The subdomain ABCD is discretized with a 40×40 cross-triangular element mesh, like the one shown in Figure 10. Traction boundary conditions corresponding to the analytical solution are applied on all boundaries other than the free surface. The stress fields σ_{xx} and σ_{xy} are compared to the exact solution in Figures 13 and 14, respectively. The convergence of the relative energy error norm is shown in Figure 12. Since the strain energy in the vicinity of the core diverges with mesh refinement, the energy within a distance of 0.05μm from the core was neglected. The convergence rate of the method is 1.0 which is the optimum rate for linear finite elements.

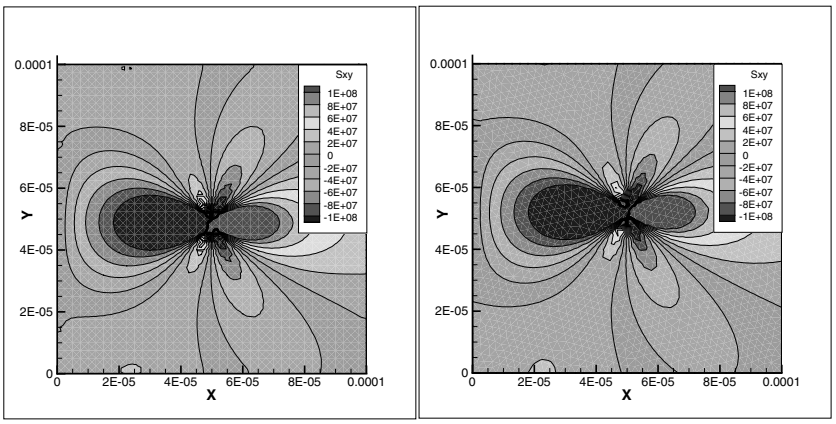


Fig. 14. Shear stress σ_{xy} in dynes/cm^2 for a dislocation in a semi-infinite domain. On the right is the exact field; on the left are the results from XFEM.

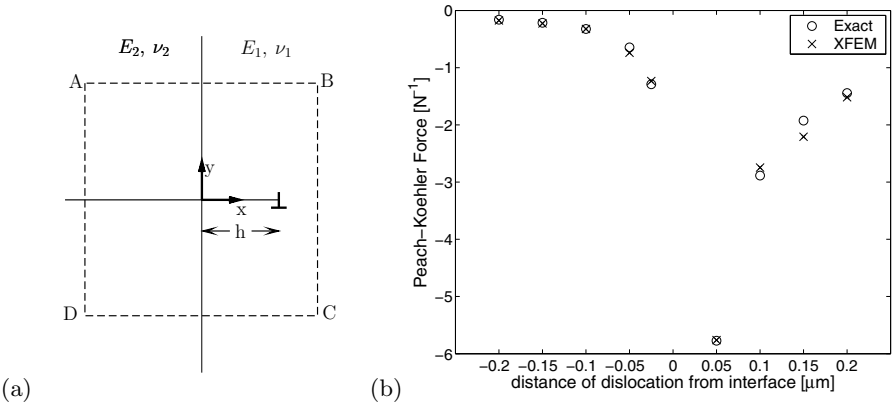


Fig. 15. (a) Nomenclature for an edge dislocation near a bimaterial interface between two semi-infinite domains. (b) comparison of the glide component of the Peach–Koehler force obtained by the proposed method with the exact result.

We consider an edge dislocation near a bimaterial interface between two semi-infinite domains, as shown in Figure 15. The bimaterial interface is located along the plane $x = 0$. An edge dislocation with Burgers vector $b = 8.5510 \times 10^{-4}$ μm with a glide plane along the plane $y = 0$ is located at $x = h$. In the subdomain $x > 0$ the elastic modulus $E_1 = 1.2141 \times 10^{11}$ Pa and Poisson's ratio $\nu_1 = 0.34$; in the subdomain $x < 0$ $E_2 = 1.2141 \times 10^{10}$ Pa and $\nu_2 = 0.3$.

The solution to this problem was first given by Head [25], and later used to study the Peach–Koehler force on a dislocation near a bimetallic interface by Dundurs et al. [26]. Further clarification of the solution was provided by Lubarda [27] in the context of dislocation arrays near bimaterial interfaces.

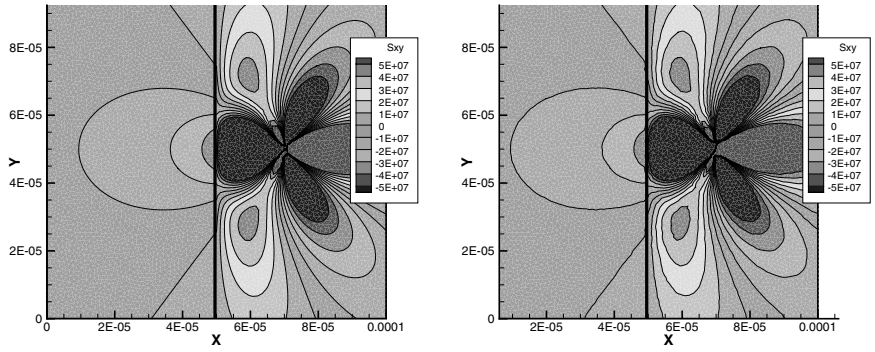


Fig. 16. Comparison of the shear stress, σ_{xy}, contours from the proposed method (right) with those of the exact solution (left) for an edge dislocation near a bimaterial interface.

For the purpose of comparison with the exact solution, we consider a square $1\mu m \times 1\mu m$ domain, centered at the origin. Along the edges of the domain, traction boundary conditions corresponding to the exact solution are applied; the expressions for the stress fields are given in [27]. We discretize the domain with an unstructured mesh of 13320 three-node triangular elements. In Figure 15, the glide component of the Peach–Koehler force obtained by a J-integral around core is compared to that from Dundurs et al. [26] for various distances, h, from the material interface.

The glide force calculated with the domain form of the J-integral compares well with the exact result. The accuracy of calculated glide force can be improved by increasing the size of the domain over which the J-integral is computed.

Contour plots of the shear stress, σ_{xy}, are shown in Figures 16 in comparison with the exact solution for $h = 0.2$ μm. The stress contours show good agreement with the exact solution fields. We see that σ_{xy} is continuous across the material interface, as expected. The far field values are especially accurate when compared to the exact solution. Further mesh refinement near the dislocation core would improve the accuracy of the stresses in the vicinity of the core.

We next illustrate the proposed method for problems with a large number of dislocations. We will consider a rectangular domain with dimensions $2 \times 10^{-4} \times 10^{-4}$ cm, elastic modulus 1.2141×10^{12} dyne/cm^2 and Poisson's ratio 0.34. A system of 200 edge dislocations with Burgers vectors $\|\mathbf{b}^\alpha\| = 8.551 \cdot 10^{-8}$ cm, $\alpha = 1, 2, ..., 200$, are considered on 14 parallel slip planes spaced $400\,\|\mathbf{b}\|$ apart with a slope of 3/2, see Figure 17. A tensial load of 1 dyne/cm^2 is applied to the right boundary. The left edge of the domain is fixed in the x-direction. Rigid body motion is constrained by fixing the node in the bottom

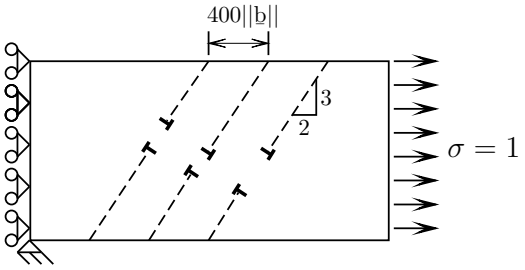


Fig. 17. Rectangular domain under uniaxial tension. Dashed lines represent glide planes.

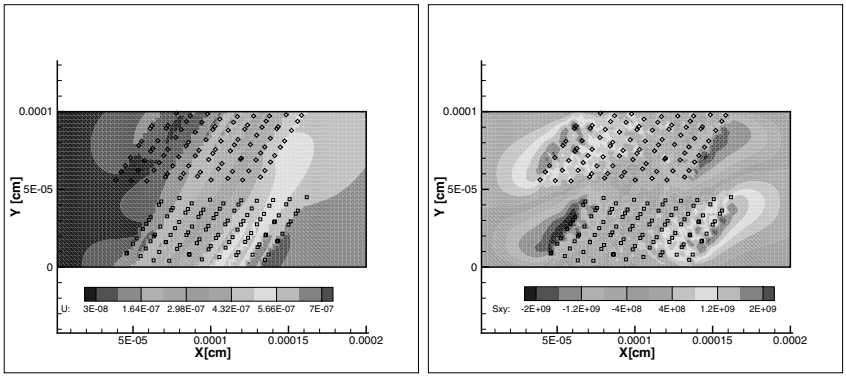


Fig. 18. Contour plots of a uniaxial tension specimen with 200 dislocations. Left is the displacement field; right is the shear stress σ_{xy} contours. Squares represent dislocation cores with positive Burgers vectors and diamonds represent dislocations with negation Burgers vectors.

left corner of the domain. The displacement and shear stress σ_{xy} contours for a uniform 80×40 element mesh are shown in Figure 18. To give a qualitative measure of the efficiency of the proposed method, the execution time for the assembly and solution of the Matlab code on a single CPU PC was about 20 seconds.

Acknowledgments

The support of the Office of Naval Research under grant N00014-98-1-0578 and the Army Research Office under grant W911NF-05-1-0049 are gratefully acknowledged.

References

1. Belytschko T,Black T (1999) International Journal for Numerical Methods in Engineering 45:601–620
2. Moës N, Dolbow J, Belytschko T (1999) International Journal for Numerical Methods in Engineering 46:131–150
3. Stolarska M, Chopp DL, Moës N, Belytschko T (2001) International Journal for Numerical Methods in Engineering 51:943–960
4. Belytschko T, Moës N, Usui S, Parimi C (2001) International Journal for Numerical Methods in Engineering 50:993–1013
5. Babuška I, Melenk JM (1997) International Journal for Numerical Methods in Engineering 40:727–758
6. Belytschko T, Chen H, Xu J, Zi G (2003) International Journal for Numerical Methods in Engineering 58:1873–1905
7. Song JH, Areias PMA, Belytschko T (2006) International Journal for Numerical Methods in Engineering 67:868–893
8. Hansbo A, Hansbo P (2004) Computer Methods in Applied Mechanics and Engineering 193:3523–3540
9. Areias PMA, Belytschko T (2006) Computer Methods in Applied Mechanics and Engineering 195:1275–1276
10. Réthoré J, Gravouil A, Combescure A (2004) Computer Methods in Applied Mechanics and Engineering 193:4493–4510
11. Réthoré J, Gravouil A, Combescure A (2005) International Journal for Numerical Methods in Engineering 63:631–659
12. Remmers JJC, de Borst R, Needleman A (2003) Computational Mechanics 31:69–77
13. Gracie R, Ventura G, Belytschko T (2007) International Journal for Numerical Methods in Engineering 69:423–441
14. Xu XP, Needleman A (1994) Journal of Mechanics & Physics of Solids 42:1397–1434
15. Camacho G, Ortiz M (1996) International Journal of Solids and Structures 33:2899–2938
16. Ortiz M, Pandolfi A (1999) International Journal for Numerical Methods in Engineering 44:1267–1282
17. Hallquist JO (1998) LS-DYNA theory manual. Livemore software thchnology corporation, USA
18. Lemaitre J, Chaboche JL (1990) Mechanics of Solid Materials. Cambridge University Press, Cambridge
19. Ramulu M, Kobayashi AS (1985) International Journal of Fracture, 27:187–201
20. Ravi-Chandar K (1998) International Journal of Fracture 90:83–102
21. Sharon E, Gross SP, Fineberg J (1995) Physical Review Letter 74:5096–5099
22. Sharon E, Fineberg J (1996) Physical Review B, 54:7128–7139
23. Fineberg J, Sharon E, Cohen G (2003) International Journal of Fracture 121:55–69
24. Hirth JP, Lothe J (1982) Theory of Dislocations: volume 1. Wiley and Sons, second edition
25. Head AK (1953) Proceedings of the Physical Society. Section B 66:793–801
26. Dundurs J, Sendeckyj GP (1965) Journal of Applied Physics 36:3353–3354
27. Lubarda VA (1997) International Journal of Solids and Structures 34:1053–1073

Some improvements of Xfem for cracked domains

E. Chahine [1], P. Laborde[2], J. Pommier[1], Y. Renard[3] and M. Salaün[4]

(1) INSA Toulouse, laboratoire MIP, CNRS UMR 5640, Complexe scientifique de Rangueil, 31077 Toulouse, France.
(2) Université P. Sabatier, laboratoire MIP, CNRS UMR 5640, 118 route de Narbonne, 31062 Toulouse cedex 4, France.
(3) INSA Lyon, Institut Camille Jordan, CNRS UMR 5208, 20, rue Albert Einstein, 69621 Villeurbanne, France.
(4) ENSICA, 1 pl. Émile Blouin, 31056 Toulouse cedex 5, France.

Summary. The XFEM method for fracture mechanics is revisited. A first improvement is considered using an enlarged fixed enriched subdomain around the crack tip and a bonding condition for the corresponding degrees of freedom. An efficient numerical integration rule is introduced for the nonsmooth enrichment functions. The lack of accuracy due to the transition layer between the enrichment area and the rest of the domain leads to consider a pointwise matching condition at the boundary of the subdomain. An optimal rate of convergence is then obtained, numerically and theoretically, even for high degree polynomial approximation.

Key words: Fracture, finite elements, XFEM, optimal rate of convergence.

1 Introduction

In many industrial situations, it is of great interest to compute the advance of a crack for a two-dimensional domain (plate, shell) or a three-dimensional body. In 1999, Moës, Dolbow and Belytschko introduced in [9] a new methodology, the so-called Xfem method, which allows to make computations in a cracked domain, the crack growth being described independently of the mesh. This is an important advantage compared to existing methods where a remeshing step and an interpolation step are necessary each time an extension of the crack is computed, which can be sources of instabilities. Our goal is to evaluate the Xfem method in terms of efficiency and quality of approximation and, as far as it is possible, to give mathematical justifications to its efficiency. We also propose some improvements for some limitations of the method.

Alain Combescure et al. (eds.), IUTAM Symposium on Discretization Methods for Evolving Discontinuities, 171–184.

2 Model Problem

A linear elastic two-dimensional problem is considered on a cracked domain Ω (see Fig. 1), with an isotropic homogeneous material. On the boundary of the non-cracked domain $\overline{\Omega}$, a Dirichlet condition is applied on Γ_D and a Neumann condition is prescribed on Γ_N. The crack is denoted by Γ_C such that Γ_D, Γ_N and Γ_C form a partition of $\partial\Omega$.

The space of admissible displacements is

$$\mathcal{V} = \{v \in H^1(\Omega; \mathbb{R}^2); v = 0 \text{ on } \Gamma_D\},$$

and the equilibrium problem is written as follows:

$$\text{find } u \in \mathcal{V} \text{ s.t. } a(u, v) = l(v) \quad \forall v \in \mathcal{V},$$

where

$$a(u, v) = \int_\Omega \sigma(u) : \varepsilon(v) \, \mathrm{d}x, \qquad l(v) = \int_\Omega f.v \, \mathrm{d}x + \int_{\Gamma_N} g.v \, d\Gamma,$$

and $\sigma(u) = \lambda \mathrm{tr}\, \varepsilon(u) I + 2\mu\varepsilon(u)$ is the stress tensor, $\varepsilon(u)$ the linearized strain tensor, $\lambda > 0$ and $\mu > 0$ the Lamé coefficients, f and g some given force densities on Ω and Γ_N respectively. A traction free condition on the crack is assumed.

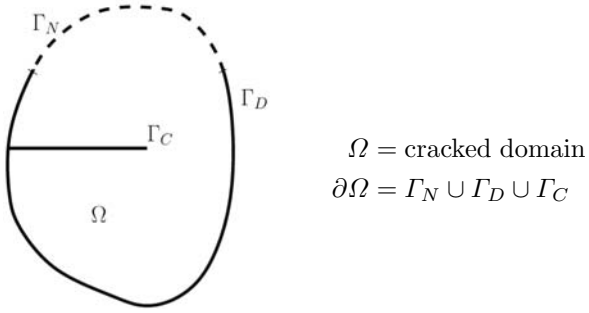


Fig. 1. A cracked domain.

The asymptotic displacement at the crack tip for the two opening modes u_I, u_{II} are given in polar coordinates (see Fig. 2) relatively to the crack tip by

$$u_I = \frac{K_I}{E}\sqrt{\frac{r}{2\pi}}\begin{pmatrix}\cos\frac{\theta}{2}\\ \sin\frac{\theta}{2}\end{pmatrix}(a + b\cos\theta)$$

$$u_{II} = \frac{K_{II}}{E}\sqrt{\frac{r}{2\pi}}(1+\nu)\begin{pmatrix}\sin\frac{\theta}{2}(c + 2 + \cos\theta)\\ \cos\frac{\theta}{2}(2 - c - \cos\theta)\end{pmatrix},$$

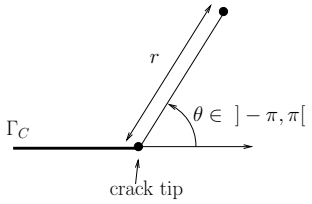


Fig. 2. Polar coordinates relatively to the crack tip.

where K_I and K_{II} are the stress intensity factors (see [8] for instance), and

$$a = 2 + \frac{2\mu}{\lambda + 2\mu}, \quad b = -2\frac{\lambda + \mu}{\lambda + 2\mu}, \quad c = \frac{\lambda + 3\mu}{\lambda + \mu}.$$

The functions u_I and u_{II} belong to $H^{3/2-\varepsilon}(\Omega; \mathbb{R}^2)$ for any $\varepsilon > 0$ (see [6], chapter 4).

3 The classical Xfem

Let X^h be a scalar finite element space on a triangular mesh of the uncracked domain $\overline{\Omega}$. For instance, X^h will represent a P_k finite element method. Let also Y^h be a P_1 finite element space on the same mesh. We will denote by φ_i the basis functions (shape functions) of X^h and by ψ_i the ones of Y^h. The classical Xfem method is then characterized by the three following techniques.

- A representation of the discontinuity of the displacement across the crack with the enrichment
$$\sum_{i \in I_H} b_i H \varphi_i,$$
(in fact $\sum_{i \in I_H} b_i H \psi_i$ was proposed in original Xfem), where $H(x)$ is the step function being equal to $+1$ on one side of the crack and -1 on the other side. The set I_H corresponds to the basis functions whose support is completely cut by the crack (see [9]). One remarks that the discontinuity is not represented on the element containing the crack tip.
- The enrichment with the asymptotic displacement at the crack tip
$$\sum_{i \in I_F} \sum_{j=1}^{4} c_{ij} F_j \psi_i,$$
where
$$F_1 = \sqrt{r}\sin\frac{\theta}{2}, \quad F_2 = \sqrt{r}\cos\frac{\theta}{2}, \quad F_3 = \sqrt{r}\sin\frac{\theta}{2}\cos\theta, \quad F_4 = \sqrt{r}\cos\frac{\theta}{2}\cos\theta,$$

and I_F is the set of basis functions whose support contains the crack tip (see [9] and Fig. 3). The singular function F_1 represents also the discontinuity on the element containing the crack tip.

- The approximation of the geometry of the crack by two level sets of functions ξ_h, ζ_h defined on a scalar fem (classically Y^h) allowing to define: $H(x) = \text{sign}(\xi_h(x)), \quad r = \sqrt{\xi_h^2 + \zeta_h^2}$, and $\theta = \arctan \frac{\xi_h}{\zeta_h}$ (see [14]).

Consequently, the finite element space approximating the displacement on the cracked domain is defined as follows:

$$\mathcal{V}^h = \left\{ v^h = \sum_i a_i \varphi_i + \sum_{i \in I_H} b_i H \varphi_i + \sum_{i \in I_F} \sum_{j=1}^{4} c_{ij} F_j \psi_i : \ a_i, b_i, c_{ij} \in \mathbb{R}^2 \right\} \quad (1)$$

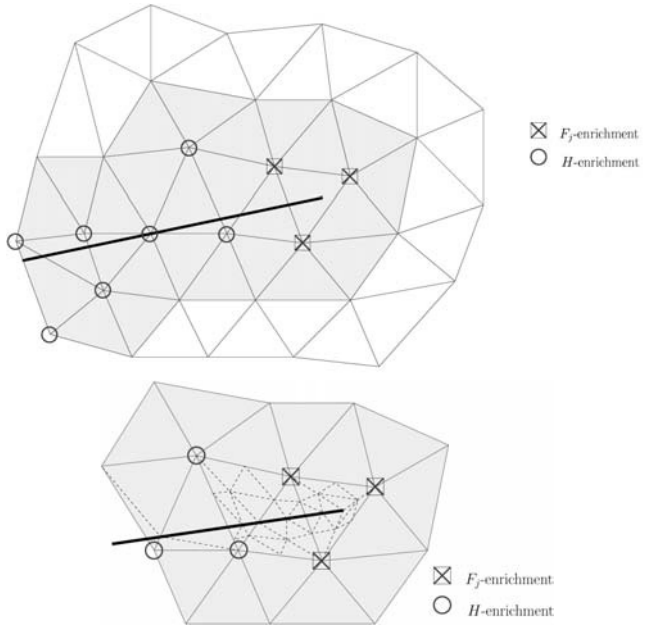


Fig. 3. Classical enrichment strategy and subdivision of the mesh for the purpose of numerical integration.

4 Representation of the discontinuity

As described in the previous section, the discontinuity of the displacement across the crack is taken into account in Xfem by the mean of an enrichment of the classical finite element space (defined on the uncracked domain) with a term of the form

$$\sum_{i \in I_H} b_i H \varphi_i,$$

where H is the step function, possibly defined via a function whose zero level set represents the geometry of the crack. It seems that this technique is the most optimal in Xfem. In [2] and [3], it is proven that the displacement is optimally approximated on each side of the crack. In our opinion, it is also one of the most original traits of Xfem. This technique can be extended to more complicated situations, for instance a situation where three zones instead of two are present, as it is illustrated in Fig. 4. This is already implemented and available in the Getfem library [11].

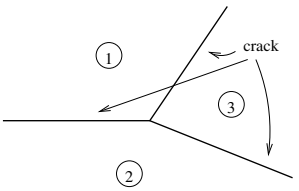


Fig. 4. A more complicated situation with three zones.

This technique can also be extended to take into account more complicated boundary conditions (for instance contact and friction condition, already treated in [5]). An important aspect to have an optimal approximation of the discontinuous part of the displacement is to use the same basis function in the enrichment as in the approximation of the displacement itself. Even when these basis functions does not represent a partition of unity (as it is the case for Hermite elements for instance). The fact that the Lagrange basis functions form a partition of unity is not the key point in that part of the enrichment.

5 Asymptotic displacement at the crack tip

In order to make a better approximation of the asymptotic displacement (and also the discontinuity) near the crack tip, a second enrichment of the form

$$\sum_{i \in I_F} \sum_{j=1}^{4} c_{ij} F_j \psi_i,$$

is considered, where I_F is the set of basis function whose support contains the crack tip. Using this strategy, the convergence curve on Fig. 5 were obtained using the first opening mode u_I as the exact solution.

Some similar results can be found in [12]. The observed rate of convergence is limited to one half, even for high order methods like P_2 and P_3. In the

following, we analyze this lack of accuracy and give some improvements for two-dimensional problems in order to recover a better convergence rate.

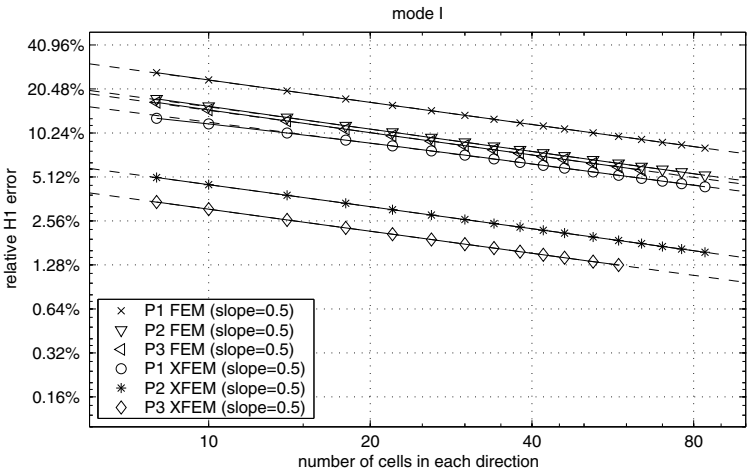
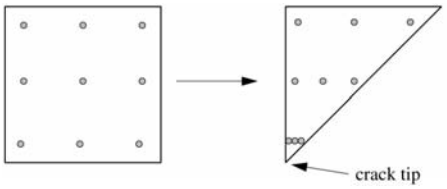


Fig. 5. Convergence curves for the classical Xfem and the first opening mode.

5.1 Adapted cubature formula for the asymptotic displacement

$$\tau : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \rightarrow \begin{pmatrix} x_1 x_2 \\ x_2 \end{pmatrix}$$



New integration points: $\overline{\xi} = \tau(\xi)$, new weights: $\overline{\eta} = \eta \det(\nabla \tau)$.

Fig. 6. The construction of adapted cubature formulas.

One of the difficulties which may degrade the quality of approximation is the fact that the gradients of the functions F_i, i = 1..4, are singular at the crack tip and the elementary integrals are not well approximated by classical cubature formulas. A remarkable aspect is that the singular part of the elementary integrals disappear when the computation is done in polar coordinates. Thus, we proposed in [7] the use of a simple adapted cubature formula on the sub-triangles having the crack tip as a vertex (see Fig. 3).

This adapted cubature formula is obtained from a classical Gauss formula on a square and using the transformation presented on Fig. 6. A convergence

test on the computation of an elementary matrix is presented on Fig. 7. Practically, with this method, 25 integration points were enough for the more accurate tests we done.

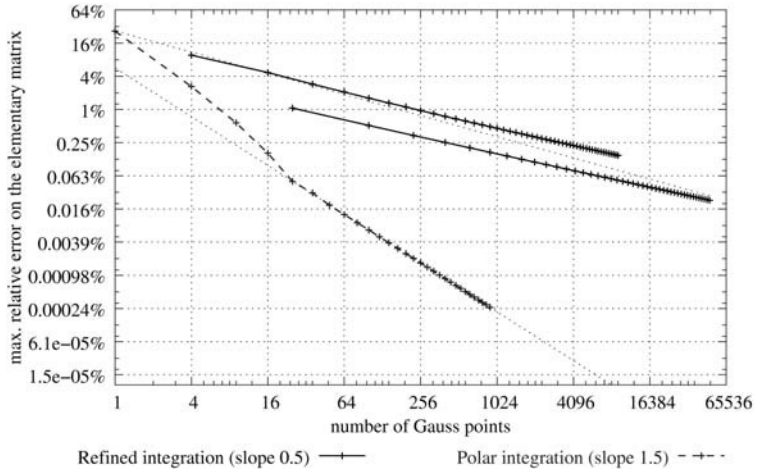


Fig. 7. Relative error on the computation of an elementary matrix with a refined order 3 Gauss method, a refined order 10 Gauss method and the almost polar method.

5.2 Fixed enriched area

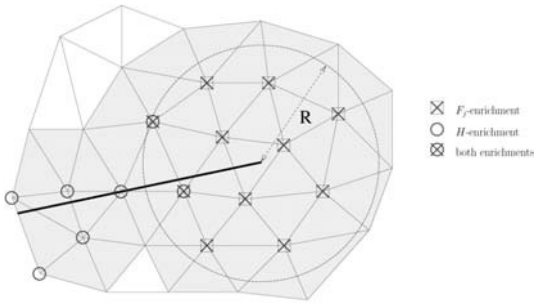


Fig. 8. Fixed enriched area method.

The second, and most important aspect which limits the convergence rate is the fact that the area enriched with the asymptotic displacement decreases with the mesh parameter h. A natural idea (independently proposed in [1]) is to have an enriched area independent of the mesh parameter. For instance, as it is presented on Fig. 8, all the nodes which are closer than a certain distance R (independant of h) will be enriched.

This gives the following enriched finite element space:

$$\mathcal{V}_R^h = \left\{ v^h \;=\; \sum_i a_i \varphi_i + \sum_{i \in I_H} b_i H \varphi_i + \sum_{i \in I_{F(R)}} \sum_{j=1}^{4} c_{ij} F_j \psi_i \;:\; a_i, b_i, c_{ij} \in \mathbb{R}^2 \right\},$$

where $I_{F(R)}$ is the set of finite element nodes contained in the disk of radius R centered on the crack-tip. The convergence curves corresponding to this method are presented on Fig. 9. The convergence rates are close to be optimal.

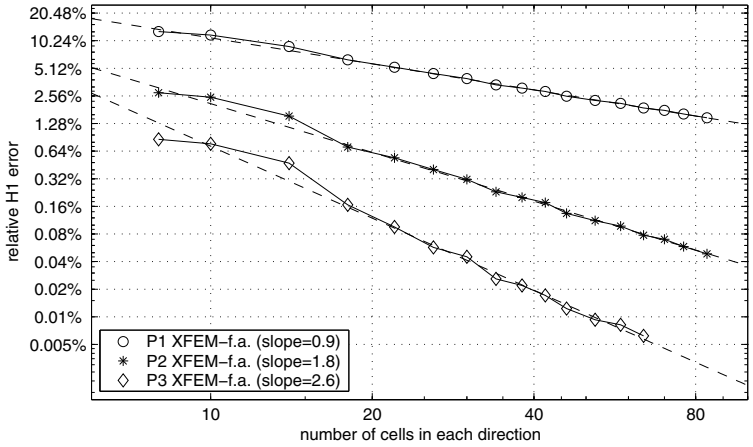


Fig. 9. Convergence curves of Xfem with a fixed enriched area.

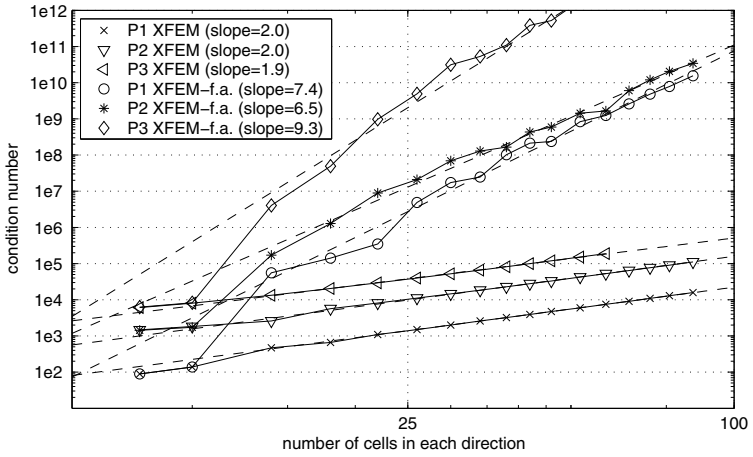


Fig. 10. Condition number of the linear system for Xfem with a fixed enriched area.

Unfortunately, as given on Fig. 10, the condition number of the linear system increases very rapidly when the mesh parameter decreases.

This is probably due to the fact that the basis functions of the enrichment are linearly dependent on an element (see [7]) and also to the fact that the functions $F_i, i = 1..4$ are somewhat flat far for the crack tip, and thus well approximated by the classical Fem.

6 Improvements for two-dimensional problems

In order to fix this difficulty, we present three techniques for two-dimensional problems:

- The dof gathering technique.
- Enrichment with the use of a cut-off function.
- Nonconforming method with a bonding condition.

All these methods are cheaper than Xfem with a fixed enriched area because the number of dof for the enrichment is very small. However, these techniques cannot be easily extended to three-dimensional problems.

6.1 The dof gathering technique

The dof gathering technique corresponds to prescribe the constraint

$$c_{ij} = c_{kj} \text{ for all } i, k \in I_{F(R)},$$

in the fixed enriched area method. This is equivalent to introducing

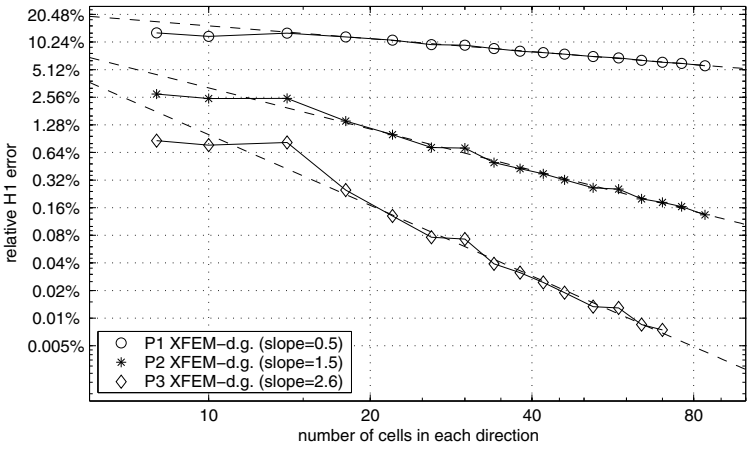


Fig. 11. Convergence curves for the dof gathering method.

$$T_R = \sum_{i \in I_{F(R)}} \psi_i,$$

and to considering the enrichment

$$\sum_{j=1}^{4} c_j F_j T_R.$$

This reduces the number of enrichment dofs to only eight (four for each of the two components, see [7]). The convergence curves of this method are presented on Fig. 11.

One can see that half an order of convergence rate is lost. After analysis, this problem is due to the transition layer between the enriched area and the rest of the domain. However, the condition number of the linear system is greatly improved by this method (see Fig. 12).

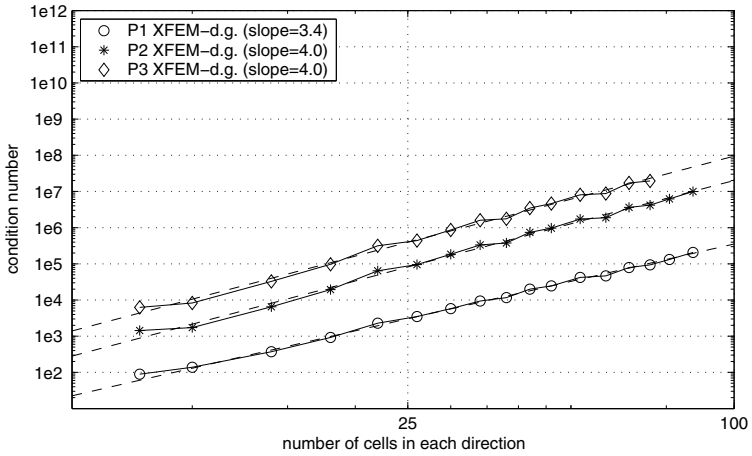


Fig. 12. Condition number of the linear system for the dof gathering method.

6.2 Enrichment with a cut-off function

Actually, the function $T_R = \sum_{i \in I_{F(R)}} \psi_i$ is a cut-off function which depends on the mesh parameter h and which is not regular. Instead of this function, the idea is to use a regular ($\mathcal{C}^1$ or $\mathcal{C}^2$) cut-off function γ which satisfies

$$\begin{cases} \gamma(r) = 1 \text{ if } r \leq r_0, \\ 0 < \gamma(r) < 1 \text{ if } r_0 < r < R, \\ \gamma(r) = 0 \text{ if } R \leq r. \end{cases}$$

where $0 < r_0 < R$. The enrichment is then defined by

$$\sum_{j=1}^{4} c_j F_j \gamma.$$

The enrichment of a finite element space by the mean of a cut-off function was already presented in [13]. The complete numerical analysis of this method in the framework of Xfem is presented in [2] and [3]. The optimal convergence rate is obtained.

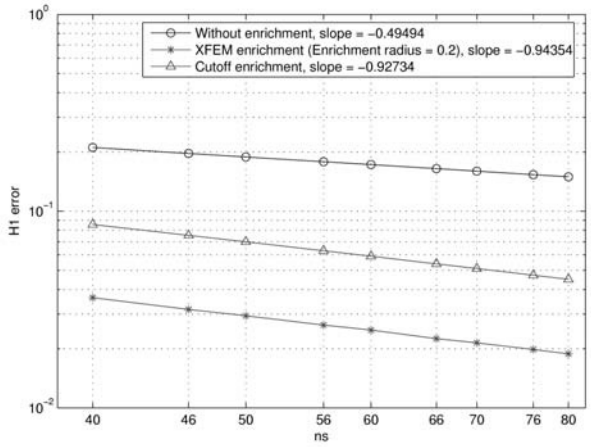


Fig. 13. Convergence curves for the cut-off method.

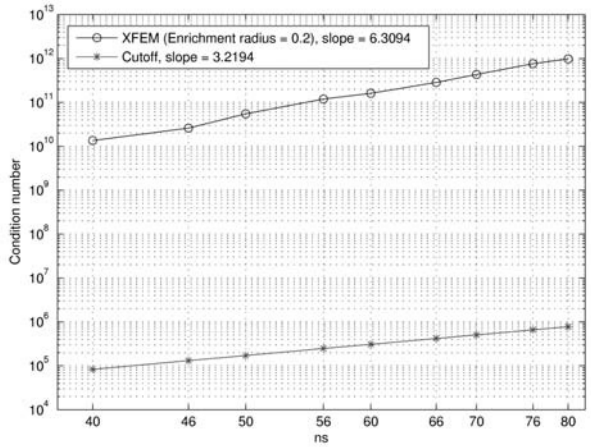


Fig. 14. Condition number of the linear system for the cut-off method.

This is illustrated on Fig. 13, where the enrichment with a cut-off function is compared to Xfem with a fixed enriched area. The optimal slope is close to be obtained. The fact that Xfem with a fixed enriched area gives a better approximation can be interpreted by the fact that the approximation is perturbated by the stiff part of the cut-off function in the present method. But, as presented on Fig. 14, the condition number is also greatly improved and the number of additional dofs is still only eight.

6.3 Enrichment with a bonding condition

Another technique, also presented in [7], is to simply enrich the finite element space by

$$\sum_{j=1}^{4} c_j F_j.$$

but only on the elements contained in the enriched area. A pointwise bonding condition is then applied on the interface Γ_I between the enriched area and the rest of the domain on each finite element node (see Fig. 15).

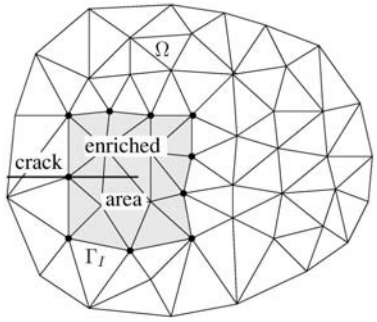


Fig. 15. Bonding condition method.

Numerically, this is the method the simplest to implement and one can see on Fig. 16 that this is also the most efficient. Fig. 17 represents a computation on a coarse mesh for the two opening modes with this technique and a P_3 method. The accuracy of the Von Mises stress is remarkable. In particular, the transition between the enriched area and the rest of the domain is not

visible.

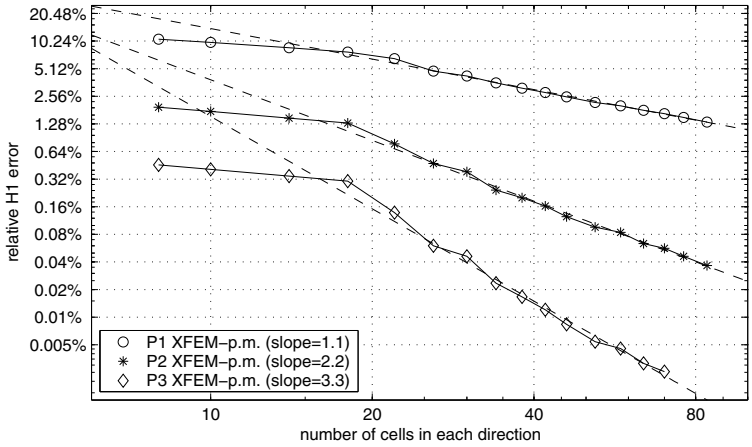


Fig. 16. Convergence curves for the enrichment with a bonding condition.

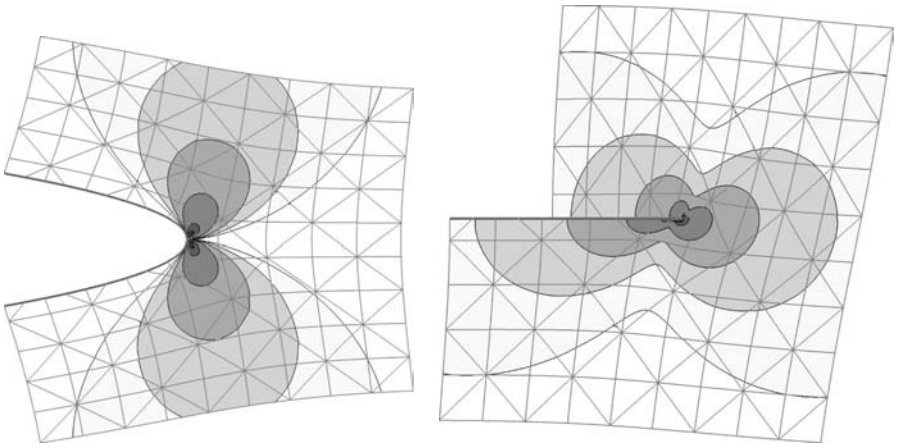


Fig. 17. P_3 XFEM solutions for the mode I and mode II problems with a pointwise bonding condition and a coarse mesh (contour levels of Von Mises stress).

Concluding remarks

In section 6 we propose three new techniques to improve Xfem method for two-dimensional domains. The extension of this work to three-dimensional problems is an open question, since the singularities are difficult to take into

account. Nevertheless the extension to plates and shells is more straightforward and is a work in progress.

References

1. E. Béchet, H. Minnebo, N. Moës and B. Burgardt. Improved implementation and robustness study of the X-FEM method for stress analysis around cracks. Submitted to *Int. J. Numer. Meth. Engng.*
2. E. Chahine, P. Laborde, Y. Renard. A quasi-optimal convergence result for fracture mechanics with XFEM. *C.R. Acad. Sci. Paris, Série I*, 342: 527–532, 2006.
3. E. Chahine, P. Laborde, Y. Renard. Crack-Tip enrichment in the Xfem method using a cut-off function. Submitted.
4. Ph. Destuynder, M. Djaoua. *Math. Meth. in the Appl. Sci.*, 3:70–87, 1981.
5. J. Dolbow, N. Moës, T. Belytschko. An extended finite element method for modeling crack growth with frictional contact. *Comput. Meth. Appl. Mech. Eng.*, 190: 6825–6846, 2001.
6. P. Grisvard. *Singularities in boundary value problems*. Masson, 1992.
7. P. Laborde, J. Pommier, Y. Renard, M. Salaün. High order extended finite element method for cracked domains. *Int. J. Numer. Meth. Engng.*, 64, 354–381, 2005.
8. J. Lemaitre, J.-L. Chaboche *Mechanics of Solid Materials*. Cambridge University Press, 1994.
9. N. Moës, J. Dolbow, and T. Belytschko. A finite element method for crack growth without remeshing. *Int. J. Numer. Meth. Engng.*, 46:131–150, 1999.
10. N. Moës, A. Gravouil and T. Belytschko. Non-planar 3D crack growth by the extended finite element and level sets, Part I: Mechanical model. *Int. J. Numer. Meth. Engng.*, 53(11):2549–2568, 2002.
11. J. Pommier, Y. Renard Getfem++. *An open source generic C++ library for finite element methods*, http://www-gmm.insa-toulouse.fr/getfem.
12. F.L. Stazi, E. Budyn, J. Chessa, and T. Belytschko. An extended finite element method with higher-order elements for curved cracks. *Computational Mechanics*, 31:38–48, 2003.
13. G. Strang, G. J. Fix. *An Analysis of the Finite Element Method*. Prentice-Hall, Englewood Cliffs, 1973.
14. M. Stolarska, D.L. Chopp, N. Moës, T. Belytschko. Modelling crack growth by level sets in the extended finite element method. *Int. J. Numer. Meth. Engng.*, 51:943–960, 2001.

2D X-FEM Simulation of Dynamic Brittle Crack Propagation

A. Combescure, A. Gravouil, H. Maigre, J. Réthoré and D. Gregoire

LaMCoS, INSA-Lyon, CNRS UMR5259, 18–20 allée des sciences, F-69621 Lyon, France, alain.combescure@insa-lyon.fr

Summary. The application of X-FEM technique to the prediction of two dimensional dynamic brittle crack growth is presented in this paper. The method is known to guarantee exact energy conservation in case of crack propagation and it is applied to the simulation of one dynamic crack propagation experiment submitted to a mixed mode loading and showing stop and restart of a crack.

Key words: X-FEM, brittle fracture, dynamic crack propagation, crack arrest, interaction integrals, cracks blunting.

1 Introduction

Modeling and predicting dynamic crack propagation remain difficult challenges. There are at least five different methods to simulate dynamic crack propagations with a finite element technology.

The first one is the element deletion strategy which consists in removing the elements which have no remaining stresses or which are fully damaged. This method, currently used in explicit codes, is very simple to implement but it has two drawbacks: very fine meshes are needed and the results exhibit mesh dependencies.

The second one is the remeshing strategy: this type of method cannot preserve energy conservation both on the whole structure and in the crack tip region.

The third one is based on "element edge" cohesive segments: this method is not robust and leads to crack paths which are mesh dependent [1] because the crack paths are forced to follow the edges of the elements.

The fourth is often named the embedded discontinuity method [2]. It allows a discontinuity of the displacement field at the element level to represent the presence of a crack. The discontinuity amplitude is typically governed by a cohesive law.

The fifth is based on X-FEM method [3]. This method may be coupled or not to a level set technique which is used to represent the geometry of the

Alain Combescure et al. (eds.), IUTAM Symposium on Discretization Methods for Evolving Discontinuities, 185–198.

crack [4]. We shall suppose here that the crack geometry is meshed with a set of 2 nodes geometrical segments. Those segments are totally independent on the structural finite element mesh. The crack growth can be either driven by a LEFM law (which can be used only for brittle fracture mechanics) or by a cohesive law approach which can rather easily be coupled with non linear behavior of the bulk material [5].

One of the main difficulties of dynamic fracture mechanics is the uncertainty upon the appropriate choice for models to be applied for dynamic crack propagation simulations and the method to obtain the corresponding material parameters. Some experiments are already available and used extensively for the validation of numerical predictions: the Kalthoff plate experiments [6] and the Rittel–Maigre ones [7, 8]. These experiments show complex crack paths but no arrest and restart. The object of this paper is first to explain a new numerical scheme based on extended finite element method and second to apply it to the simulation of a very carefully designed experiment which provides precise and reliable data on dynamic brittle crack propagations. The paper is limited to bi-dimensional experiments and analysis. The numerical simulation of dynamic crack growth is a useful tool to understand the physics of the dynamic crack propagation. For instance it is difficult to explain from the measures alone the reasons why the crack changes direction or stops during the propagation. The numerical simulation helps to understand such complex questions: for instance the stop and direction changes can be explained by stresses wave reflections during the transient event, which are very difficult to obtain experimentally. The paper is organized as follows:

The concepts used for brittle fracture mechanics numerical simulation are first presented.

The use of the X-FEM tool for dynamically evolving cracks is presented in a second section.

The experiment and its numerical simulation are then presented.

2 Dynamic brittle fracture mechanics

The dynamic crack propagation is a rather difficult subject because the experimental check of the theoretical models is difficult. We shall limit the thermodynamical presentation to a bi-dimensional case and only one crack of length a.

2.1 Theoretical background

Let us suppose that the crack propagates at a velocity $\dot{a}$, the total energy variation writes:

$$\frac{dW}{dt} = \left(\frac{\partial W}{\partial t}\right)_a + \left(\frac{\partial W}{\partial a}\right)_t \dot{a}. \tag{1}$$

In equation (1) W is the total energy which includes kinetic and external loads energy. The second term of equation (1) has a dissipative nature and represents the energy necessary to advance the crack. A detailed and precise theory can be found in [9, 10]. The definition of energy release rate G is the following in case of a load which is not applied on the crack lips:

$$G = \lim_{\Gamma \to 0} \frac{F}{\dot{a}} = \frac{1}{\dot{a}} \lim_{\Gamma \to 0} \left[\int_{\Gamma} ((w + \tau)\dot{a} n_1 + \sigma_n \cdot \dot{u}) d\Gamma \right]. \tag{2}$$

In equation (2) Γ is a closed contour surrounding the crack tip, w (resp. τ) the strain (resp. kinetic) energy density, n_1 the normal to the contour Γ, σ_n the traction vector normal to the contour Γ and $\dot{u}$ the velocity field. This definition can be used for any type of dynamic crack advance and any type of material. Under the hypothesis of linear material behavior and stationary crack tip fields the G integral can be proved to be independent of Γ [9, 10] and its expression is:

$$G = \int_{\Gamma} ((w + \tau)\delta_{1j} + \sigma_{ij} u_{i,1}) n_j d\Gamma. \tag{3}$$

In equation (3) the convention of summation on repeated indices is supposed, δ_{1j} is the Kronecker symbol, 1 is the direction of the crack tip and n_j the outward jth component of the normal to Γ contour. Let us now define the following stress intensity factors:

$$\begin{cases} K_I^{\text{kin}} = \lim\limits_{r \to 0} \dfrac{\mu}{4(1-\nu)} \sqrt{\dfrac{2\pi}{r}} \, [u_2(\theta = \pi)], \; K_I^{\text{dyn}} = \lim\limits_{r \to 0} \sqrt{2\pi r} \, \sigma_{22}(\theta = 0), \\ K_{II}^{\text{kin}} = \lim\limits_{r \to 0} \dfrac{\mu}{4(1-\nu)} \sqrt{\dfrac{2\pi}{r}} \, [u_1(\theta = \pi)], \; K_{II}^{\text{dyn}} = \lim\limits_{r \to 0} \sqrt{2\pi r} \, \sigma_{12}(\theta = 0). \end{cases} \tag{4}$$

In equation (4), 1 denotes the direction of the tangent direction of the crack at the crack tip, ν the Poisson's ratio and μ the second Lame coefficient. The intensity factors are related through the following relations:

$$K_i^{\text{kin}} = f_i(\dot{a}) K_i^{\text{dyn}} \quad i \in \{I, II\}, \tag{5}$$

where

$$\begin{cases} f_i(\dot{a}) = \dfrac{4\alpha_i(1 - \alpha_{II}^2)}{(\kappa + 1) D(\dot{a})}, \\ \alpha_i = \sqrt{1 - \left(\dfrac{\dot{a}}{c_i} \right)^2} \\ D(\dot{a}) = 4\alpha_I \alpha_{II} - (1 + \alpha_{II}^2)^2. \end{cases} \tag{6}$$

In equation (6), the indices i are I or II, κ is the Kolosov constant (1 for plane stress state and 3–4ν for plane strain case), c_I is the compression elastic wave velocity and c_{II} the shear wave one. The zero of D function gives the

Rayleigh wave velocity which will be denoted c_R. The two f_i functions have a value of 1 for a null crack speed. Injecting equations (4), (5), (6) in equation (3) one gets:

$$G = \frac{1}{E^*}\left(K_I^{\text{kin}} K_I^{\text{dyn}} + K_{II}^{\text{kin}} K_{II}^{\text{dyn}}\right), \tag{7}$$

$$E^* = \begin{cases} E & \text{plane stress,} \\ \dfrac{E}{1-\nu^2} & \text{plane strain.} \end{cases} \tag{8}$$

These equations are the extension of Irwin's equation [11] to the dynamic case. If we now want to predict dynamic crack propagation, we have to answer three questions:

1. In which direction will the crack propagate?
2. At what speed will it run?
3. Will the crack propagate?

We shall now limit the presentation to the case where the crack propagation direction is driven by the maximum principal hoop stress. The direction is then given by the following equation:

$$\theta_c = 2\arctan\left[\frac{1}{4}\left(\frac{K_I^{\text{dyn}}}{K_{II}^{\text{dyn}}} - \text{sign}(K_{II}^{\text{dyn}})\sqrt{8 + \left(\frac{K_I^{\text{dyn}}}{K_{II}^{\text{dyn}}}\right)^2}\right)\right]. \tag{9}$$

Knowing the possible critical direction θ_c, one defines an equivalent stress intensity factor K_I^* by the following equation:

$$K_I^* = \cos^3\left(\frac{\theta_c}{2}\right) K_I^{\text{dyn}} - \frac{3}{2}\cos\left(\frac{\theta_c}{2}\right)\sin(\theta_c) K_{II}^{\text{dyn}}. \tag{10}$$

The crack is supposed to propagate at the speed which is such that:

$$\begin{cases} \dot{a} > 0 \text{ with } K_I^* = K_{I_D}(\dot{a}), \\ \dot{a} = 0 \text{ if } K_I^* < K_{I_C}, \\ K_{I_D}(\dot{a}) = \dfrac{K_{I_C}}{1 - \frac{\dot{a}}{c_R}}. \end{cases} \tag{11}$$

In equation (11) K_{I_C} is the static fracture toughness. Equation (11) is a nonlinear equation when the crack speed is not zero. These equations are sufficient to define brittle fracture crack propagation.

2.2 Numerical evaluation of fracture parameters

The precise numerical evaluation of the stress intensity factors is a rather difficult task. The line integral evaluation given in equation (3) is not easy to compute with finite elements because the best values of stresses are obtained on gauss points within the elements. The best numerical method consists to replace the line integral by a surface integral using the divergence theorem. The evaluation of dynamic stress intensity factors rely on the interaction integral presented in the book of Freund [10] and denoted M integral by Attigui [12]. This integral writes:

$$I^{\text{int}} = -\int_S q_{i,j}[(\sigma_{kl}^{\text{aux}} u_{k,l} - \rho \dot{u}_k \dot{u}_k^{\text{aux}})\delta_{ij} - (\sigma_{kj}^{\text{aux}} u_{k,i} + \sigma_{kj} u_{k,i}^{\text{aux}}]dS$$
$$+ \int_S q_i[(\sigma_{kj,j}^{\text{aux}} u_{k,i} + \rho \ddot{u}_k u_{k,i}^{\text{aux}}) + \rho(\dot{u}_k \dot{u}_{k,i}^{\text{aux}} + \dot{u}_k^{\text{aux}} \dot{u}_{k,i})]dS. \quad (12)$$

In this equation S is the surface surrounded by the contour Γ, q is a virtual extension field which must be continuous and tangent to the crack lips and zero out of the Γ contour, $(u^{\text{aux}}, \dot{u}^{\text{aux}}, \sigma^{\text{aux}})$ are auxiliary displacement velocity and stress fields which are arbitrary but must be statically and kinematically admissible. Figure 1 shows a typical virtual extension field q and also displays the integration box around the crack tip.

If one chooses the q field module to be 1 at crack tip one has:

$$I^{\text{int}} = \frac{2}{E^*}(f_I(\dot{a})K_I^{\text{dyn}} K_I^{\text{aux}} + f_{II}(\dot{a})K_{II}^{\text{dyn}} K_{II}^{\text{aux}}). \quad (13)$$

It is now clear that if one chooses the auxiliary fields to be such that $K_I^{\text{aux}} = 1$, $K_{II}^{\text{aux}} = 0$ one will get directly K_I^{dyn} from the interaction integral. The same holds for the computation of K_{II}^{dyn}. Let us now compute the dynamic stress intensity factors: the simplest solution is to choose the analytical mode I and mode II solutions as auxiliary stress fields. The computation of these quantities implies the evaluation of integrals which are performed numerically using a domain S defined by a rectangular box filled with a large number of sub-elements using standard Gauss integration schemes. The auxiliary fields are computed using the analytical expressions and the other field values within the sub-element are calculated using the original mesh with its associated shape functions.

2.3 Remarks on extension to other cases like different brittle failure mechanism or ductile dynamic fracture

The preceding formulation is only valid for brittle fracture mechanics because of the use of energy release rate G concept. If one would like to use this type of approach for a different critical direction one has to replace equations (9) and (10) by different ones. For ductile fracture one has to come back to the

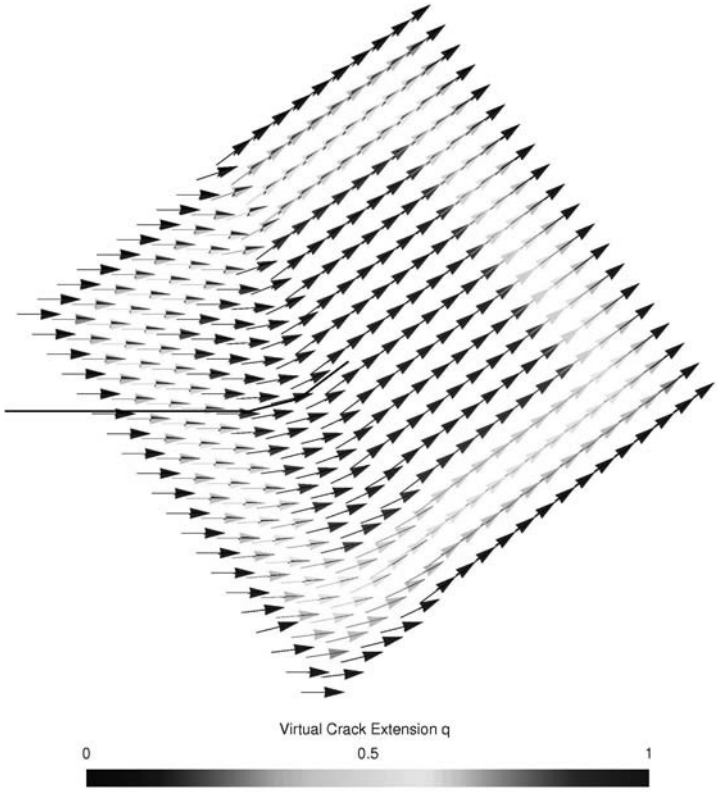


Fig. 1. Virtual crack extension field q and associated integration box.

basic definition given by equation (2) and hence to use the stress state close to the crack tip to define the properties of the propagation as it is often used in the cohesive zone approach of the simulation [13]. One may nevertheless observe that in existing cohesive zone models the velocity of the crack tip is governed by the de-cohesion rate.

3 X-FEM method for dynamically propagating cracks

3.1 X-FEM strategy: description and energy conservation properties

We shall describe in this section the special features associated with X-FEM simulation of dynamically evolving brittle cracks. The standard X-FEM method is used here [3] with the usual jump function added into the elements fully cut by the crack and the singular functions for the element partially

cut. The following representation of displacement field within the elements is chosen:

$$u(\xi,n) = \sum_{i=1}^{\text{nodes}} N_i(\xi,\eta)[u_i + H_i(\xi,\eta)a_i] \quad \text{for cut elements,}$$

$$u(\xi,n) = \sum_{i=1}^{\text{nodes}} N_i(\xi,\eta)\left[u_i + \sum_{j=1}^{4} \psi_{ij}(\xi,\eta)b_{ij}\right] \quad \text{crack tip element.} \quad (14)$$

In equation (14) $N_i(\xi,\eta)$ are the standard shape functions, a_i represent the amplitude of the degrees of freedom associated with the jump function H and b_{ij} represent the amplitude of the jth singular functions at node i of the element containing the crack tip. Let us recall that the following 4 singular ψ_{ij} functions are chosen:

$$\sqrt{r}\left\{\sin\left(\frac{\theta}{2}\right), \cos\left(\frac{\theta}{2}\right), \sin\left(\frac{\theta}{2}\right)\sin(\theta), \cos\left(\frac{\theta}{2}\right)\sin(\theta)\right\}. \quad (15)$$

It must be observed that these extended functions are implicitly linked to the crack tip position (r is the distance to the crack tip and θ the angle with the tangent to the crack tip direction). When the crack propagates statically in an elastic medium, one can forget the previous singular function for the new position of the crack. This is not possible in time dependent problems, such as dynamic crack propagations or elasto-viscoplastic material behavior. For dynamic crack propagation it has been shown [14] that if one simply adds new extended degree of freedom associated with the new crack position and that these new extended displacements and velocities are initialized to 0. Consequently the energy is perfectly kept and the new crack segment is closed at the beginning of the time step. Within the next time step the crack opens and the opening work is exactly equal to the fracture energy necessary to propagate the crack. This is a very important result because it ensures that the change of geometry of the body (crack length change) does not affect the numerical results. The local energy is not altered by the geometry changes.

3.2 Time integration

The simplest implementation is to use a standard Newmark scheme to integrate the transient equations. The implicit integration is straight forward. One could prefer to use an explicit integrator because the crack propagation is generally very fast. But the method has to be adapted because the time step tends to zero when the crack passes very close to a node. This is simply due to the fact that the mass of the jump degree of freedom of the node very close to the interface is close to zero. A simple approximation has been proposed in [15] to overcome this difficulty: one replaces the added mass on the

jump degrees of freedom (a_i) by the standard translational mass for the discontinuous enrichment. The mass matrix is diagonal and the time step of the explicit extended finite element constructed by this method has been shown analytically to be greater than one half of the standard time step of the non enriched element: this is valid for linear finite elements.

3.3 Specific crack propagation strategy

The crack propagation evaluation is a complex task because the equations defining the direction, the equivalent stress intensity and the crack speed are coupled and non linear (equations (9), (10), (11), (13)). The simplest way to proceed is to use an explicit time integration of crack advance with the known quantities at the beginning of the time step. In case of propagation the time step is generally chosen to be such that the crack does not progress of more than one element. Nevertheless the time step has to be large enough for the crack tip not to stay more than two time steps within a single element (because of the enrichment strategy choice that keeps the enriched functions corresponding to the old position of the tip). In case of dynamic explicit computation this leads to a constraint for the time steps which may not be compatible with the CFL condition. This explicit procedure works but sometimes induces instabilities in the crack propagation simulation. It needs to be improved in order to permit robust computations: a possible improvement is the implicitation of the prediction of crack increment.

4 Numerical interpretation of experimental test

This section describes a specific experiment for dynamic crack propagation. It uses a Hopkinson's bar system, and can be considered as a reference case. This experiment shows very interesting features as complex crack paths and crack arrest and restart.

4.1 Experimental set up

The experimental loading system is displayed in Figure 2.

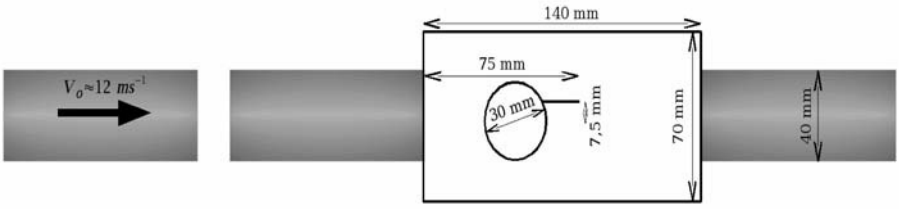


Fig. 2. Hopkinson's bar characteristics.

Table 1. Material data for the specimen (PMMA) and bars (Nylon).

Property name	Unit	PMMA	Nylon
Young's Modulus	GPa	3.3	3.6
Poisson's ratio	–	0.42	0.41
Density	Kg/m^3	1180	1145
Rayleigh wave speed	m/s	938	996

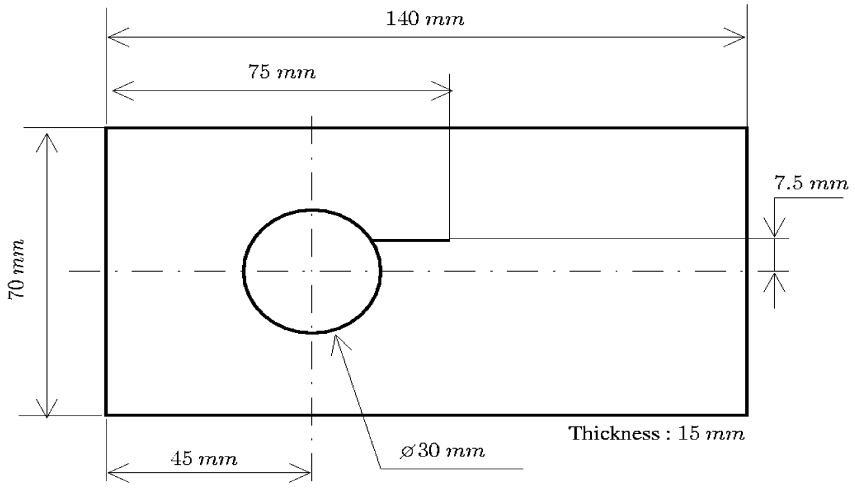


Fig. 3. Detailed geometry of the test specimen.

The specimen is a simple rectangular PMMA plate with a wave concentrator (the circular hole) and an initial non symmetric crack. It is placed between two nylon bars with appropriate sections. The sections are close to that of the PMMA specimen (1050 mm^2). This permits to avoid the wave reflection on the interface. The thickness of the plates is 15 mm. Careful experiments are performed: they are found to be very repetitive, which proves the good quality of the loading system. The input and output velocities as well as the input and output stress waves have been measured and treated in order to have usable data for the numerical analysis.

4.2 Material data

Table 1 describes the basic material data. The determination of Young's modulus for a PMMA under dynamic load is a difficult task because the elastic properties of this material are strain rate depend- ant. A mean value is proposed here. This value is chosen in order to get the best possible fit between experimental results and simple one dimensional wave propagation model of the experiment.

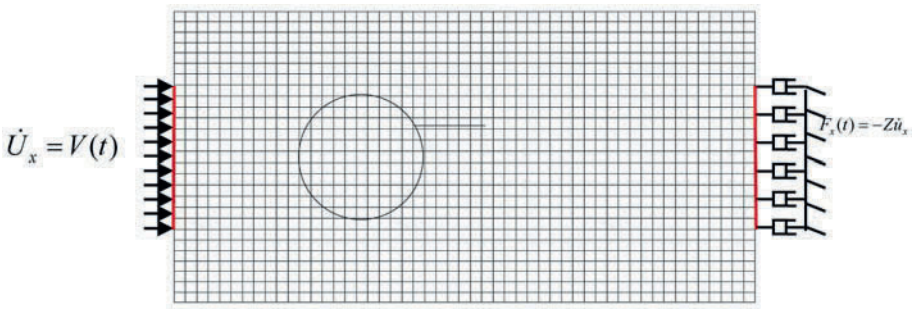


Fig. 4. Mesh of the specimen.

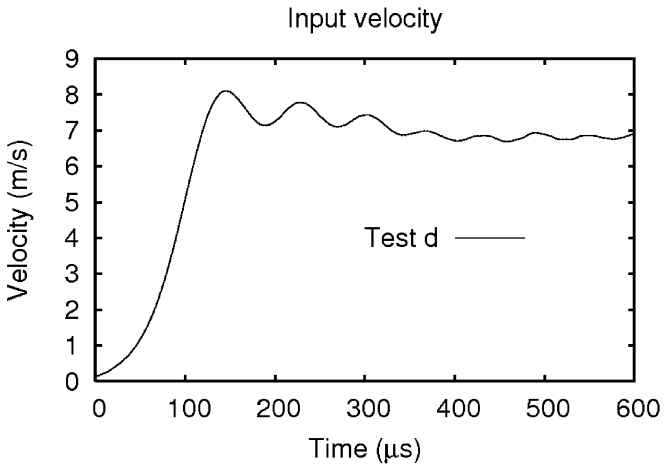


Fig. 5. Input velocity time history.

4.3 One hole specimen

The precise geometry is displayed in Figure 3.

The first modeling difficulty is the specification of appropriate boundary conditions. One could mesh the two nylon bars, but these bars are axisymetric and very long to avoid wave reflection. The resulting model would lead to axisymetric plane strain coupling, as well as a large model which would lead to heavy computations. A plane strain model of the specimen alone was used. The right hand side (output) bar is modeled by an impedance condition. This condition is applied on all nodes on a width of 40 mm. The impedance is a linear relationship between the velocity of the interface and the applied force on the interface. The following equation gives the impedance relationship:

$$F_x(t) = -\sqrt{E_{\text{nylon}}\rho_{\text{nylon}}}\,\dot{u}_x. \tag{16}$$

The second difficulty is the modeling of the moving crack path. Had the crack's path been known, an "element edge" cohesive zone approach could have been

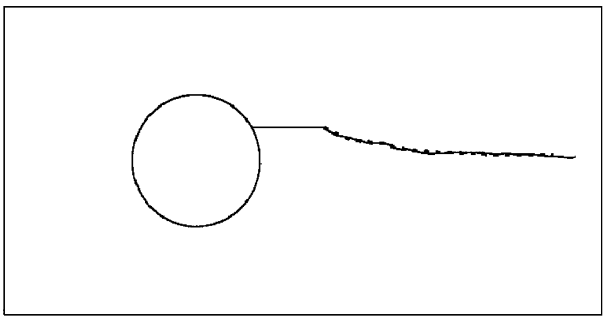


Fig. 6. Comparison of measured (dotted line) and experimental (continuous line) crack paths ($K_{I_C} = 133$ MPa$\sqrt{\text{m}}$).

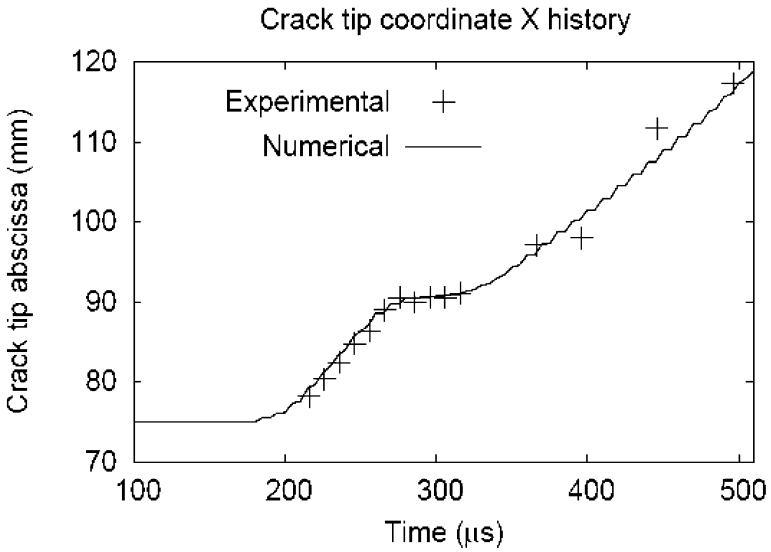


Fig. 7. Comparison of computed and experimental horizontal crack speed ($K_{I_C} = 1.33$ MPa$\sqrt{\text{m}}$).

successful, but it cannot be used in order to predict the crack's path prior to the experiment. The mesh, consisting of 1,500 4-node elements, is shown in Figure 4. The number of enriched X-FEM element is changing with time. The loading of the left impacting bar is simulated by an imposed speed on the left of the model.

The measured velocity $V(t)$ is applied to the left part of the specimen, on a 40 mm width. The input velocity time history is given by Figure 5.

The computation has been performed with a constant value of K_{I_C} (1.33 MPa$\sqrt{\text{m}}$). The comparison of computed and experimental crack path is displayed in Figure 6. The comparison of crack velocity is given in Figure 7.

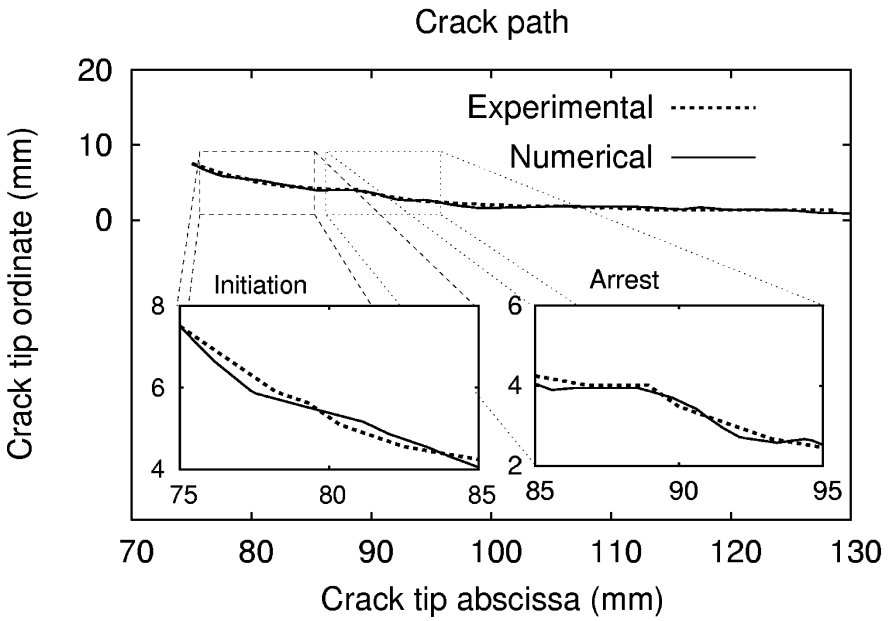


Fig. 8. Comparison of computed and experimental crack path at initiation of crack propagation ($K_{I_C} = 1.33$ MPa$\sqrt{\text{m}}$).

The agreement is rather good except at the very beginning of the propagation: the crack path is slightly different from the measured one (Figure 8). An explanation of this difference could be the radius of curvature of the notch, which is really smaller when the crack propagates.

Figure 9 shows a photograph of the crack tip just after propagation. It is clear that the COD of the moving crack is much smaller of that of the arrested one. The computation has then been rerun with two different values of K_{I_C}. When the crack is at arrest ($\dot{a} = 0$) the value is taken as the static value of 1.42 MPa$\sqrt{\text{m}}$ whereas the previous value is kept for running cracks (1.33 MPa$\sqrt{\text{m}}$). With these values the trajectories as well as crack velocities perfectly coincide with experimental data.

5 Conclusion

This paper has shown the interest of X-FEM to model dynamic brittle crack propagation. Simple macroscopic models seem to be sufficient to explain the crack complex propagations. Indeed, the results of an experimental dynamic brittle crack growth are compared successfully with numerical X-FEM simulation.

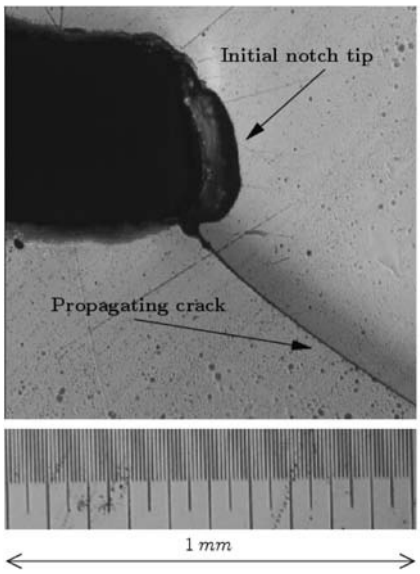


Fig. 9. Photograph of propagating crack tip.

Acknowledgements

The LMS joint laboratory between CNRS and Ecole Polytechnique (France) has provided the experimental test rig.

References

1. Ortiz M, Pandolfi A (1999) International Journal for Numerical Methods in Engineering, 44, 277–282
2. Jirásek M (2000) Computer Methods in Applied Mechanics and Engineering, 188, 307–330
3. Moes N, Dolbow J, Belytschko T (1999) International Journal for Numerical Methods in Engineering, 46(1), 133–150
4. Gravouil A, Moes N, Belytschko T (2002) International Journal for Numerical Methods in Engineering, 53(11), 2549–2586
5. de Borst R, Remmers J, Needleman A (2006) Engineering Fracture Mechanics, 76(2), 160–177
6. Kalthoff J (2000) International Journal of Fracture, 101, 1–31
7. Rittel D, Maigre H (1996) Mechanics of Materials, 23, 229–239
8. Rittel D, Maigre H (1996) Mechanics Research Communications, 23, 475–481
9. Bui H D (1978) Mécanique de la rupture fragile, Paris, Masson
10. Freund L (1990) Dynamic fracture mechanics, Cambridge, Cambridge University Press
11. Irwin G (1957) Journal of Applied Mechanics, 24(3), 361–364
12. Attigui M, Petit C (1997) International Journal of Fracture, 84, 19–36

13. de Borst R, Remmers J, Needleman A, Abellan M (2004) International Journal for Numerical and Analytical Methods in Geomechanics, 28, 583–607
14. Réthoré J, Gravouil A, Combescure A (2005) International Journal for Numerical Methods in Engineering, 63, 631–659
15. Menouillard T, Réthoré J, Combescure A, Bung H (2006) International Journal for Numerical Methods in Engineering, 68(9), 911–939

A numerical framework to model 3-D fracture in bone tissue with application to failure of the proximal femur

T. Christian Gasser[1] and Gerhard A. Holzapfel[1,2]

[1] Royal Institute of Technology (KTH), Department of Solid Mechanics
Osquars Backe 1, SE–100 44 Stockholm, Sweden
[2] Graz University of Technology, Computational Biomechanics
Schiesstattgasse 14-B, A–8010 Graz, Austria
gh@hallf.kth.se

Summary. Bone can be regarded as a quasi-brittle material. Under excessive loading nonlinear fracture zones may occur ahead the crack tips, where, typically, cohesive mechanisms are activated. The finite element method provides a powerful tool to analyze fracture formations on a numerical basis, and to better understand failure mechanisms within complex structures. The present work aims to introduce a particular numerical framework to investigate bone failure. We combine the partition of unity finite element method with the cohesive crack concept, and a two-step predictor-corrector algorithm for tracking 3-D non-interacting crack paths. This approach renders a numerically efficient tool that is able to capture the strong discontinuity kinematics in an accurate way. The prediction of failure propagation in the proximal part of the femur under compressive load demonstrates the suitability of the proposed concept. A 3-D finite element model, which accounts for inhomogeneous fracture properties, was used for the prediction of the 3-D crack surface. The achieved computational results were compared with experimental data available in the literature.

Key words: Bone, Cohesive zone model, Crack Initialization, Crack tracking algorithm, Partition of unity finite element method, Strong discontinuity, Variational formulation.

1 Introduction

Proximal femur fracture represents an important public health-care problem. Fracture of the proximal part of the femur is frequently observed in the elderly population, and it most commonly arises from accidental fall. Because of its surgical complexity and its socio-economical importance this type of bone failure is under extensive scientific investigation, and a large number

Alain Combescure et al. (eds.), IUTAM Symposium on Discretization Methods for Evolving Discontinuities, 199–211.

of 3-D finite element computations have been conducted in the past, see, for example, [14, 15, 6, 22, 13, 11]. One main focus in bone mechanics is the utilization of different fracture criteria within the finite element method. Hence, several researchers focus attention on the failure initialization, and not on the prediction of failure propagation. To the authors' knowledge, solely the work by Ota et al. [21] deals with a finite element fracture analysis of the proximal part of the femoral bone with the goal to estimate the onset and the progress of fracture in a patient-specific 3-D model. Although it can be seen as a pioneering study, it does, however, not say much about the numerical concept applied, and the underlying fracture mechanical assumptions.

Bone tissue consists mostly of collagen and mineral in the form of hydroxyapatite crystals [4]. The heterogeneous structure of bone tissue (over several length scales) causes nonlinear fracture zones ahead the tip, where, typically, cohesive mechanisms are activated. In particular, collagen fiber bridging causes nonlinear cohesive zones which are in the millimeter range, and hence linear elastic fracture mechanics does not apply anymore [29], [30].

The present work aims to model the onset and the progress of fracture in the proximal femur under compressive loading. We pursue here the strong discontinuity approach, and the existence of a fracture process zone is postulated. Hence, bone fracture is regarded as a gradual process in which separation between incipient material surfaces is resisted by cohesive traction. We assume mode-I failure properties are applicable to characterize the mixed mode situation present, and employ an isotropic (discrete) constitutive description of the cohesive zone introduced in [8]. The proposed numerical failure model combines the cohesive crack concept with the **P**artition of **U**nity **F**inite **E**lement **M**ethod (PUFEM), which has been shown to be an effective and robust numerical method to study crack propagation in concrete [28], [8], and in arterial tissue [10]. Most important for the application of this concept to patient-specific bone models is a 3-D crack tracking algorithm. Within this study we used the (partially local two-step) predictor-corrector algorithm [9], which provides the required geometrical information to handle multiple non-interacting cracks.

In Section 2 the applied modeling assumption, i.e. the underlying continuum mechanical basis (kinematics, variational formulation and constitutive models), the applied finite element implementation (PUFEM, crack initialization criteria and the crack tracking algorithm) are briefly discussed. The proposed femoral fracture model and its predictions are discussed in detail in Section 3, while in Section 4 conclusions are drawn and the work's limitations are discussed.

2 Modeling assumptions

In this section we introduce the applied modeling assumptions. In particular, the continuum mechanical basis and the proposed finite element implementa-

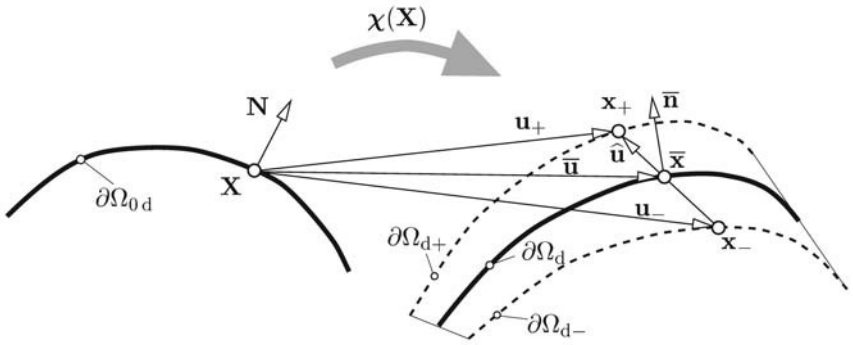


Fig. 1. Strong discontinuity kinematics capturing bone fracture.

tion are addressed. A more comprehensive description and derivation of the underlying numerical concept is given in [8], [9], and references therein.

2.1 Continuum mechanical basis

Strong discontinuity kinematics

We pursue the strong discontinuity approach and assume a discontinuity in the displacement field that characterizes bone failure. In particular, it is assumed that a discontinuity $\partial\Omega_{0\,\mathrm{d}}$ separates the body $\partial\Omega_0$ into two sub-bodies[3], which themselves occupy the referential sub-domains Ω_{0+} and Ω_{0-}, so that $\partial\Omega_{0\,\mathrm{d}} \cap \Omega_{0+} \cap \Omega_{0-} = \emptyset$ and $\Omega_{0+} \cup \partial\Omega_{0\,\mathrm{d}} \cup \Omega_{0-} = \Omega_0$. A deformation $\boldsymbol{\chi}(\mathbf{X})$ maps Ω_{0+} and Ω_{0-} into their (related) current configurations Ω_+ and Ω_-. Consequently, a displacement at a material point $\mathbf{X}$ is assumed to be $\mathbf{u}(\mathbf{X}) = \mathbf{u}_{\mathrm{c}}(\mathbf{X}) + \mathcal{H}(\mathbf{X})\mathbf{u}_{\mathrm{e}}(\mathbf{X})$, where $\mathbf{u}_{\mathrm{c}}$ and $\mathbf{u}_{\mathrm{e}}$ are regular and enhanced displacement fields, respectively [18], [2]. Here $\mathcal{H}(\mathbf{X})$ denotes the *Heaviside* function, with the values 0 and 1 for $\mathbf{X} \in \Omega_{0-}$ and $\mathbf{X} \in \Omega_{0+}$, respectively. In addition, we introduce the unit normal vector $\mathbf{N}(\mathbf{X})$, which defines the orientation of $\partial\Omega_{0\,\mathrm{d}}$ in $\mathbf{X}$, see Fig. 1.

Standard derivation [17], [12] and the use of $\mathrm{Grad}\,\mathcal{H}(\mathbf{X}) = \delta_{\mathrm{d}}\mathbf{N}(\mathbf{X}_{\mathrm{d}})$ defines the corresponding deformation gradient

$$\begin{aligned}\mathbf{F}(\mathbf{X}) &= \mathbf{I} + \mathrm{Grad}\,\mathbf{u}(\mathbf{X}) \\ &= \mathbf{I} + \mathrm{Grad}\,\mathbf{u}_{\mathrm{c}}(\mathbf{X}) + \mathcal{H}\mathrm{Grad}\,\mathbf{u}_{\mathrm{e}}(\mathbf{X}) + \delta_{\mathrm{d}}(\mathbf{X})\mathbf{u}_{\mathrm{e}}(\mathbf{X}) \otimes \mathbf{N}(\mathbf{X}_{\mathrm{d}}), \quad (1)\end{aligned}$$

which serves as the basis for the right Cauchy-Green tensor $\mathbf{C} = \mathbf{F}^{\mathrm{T}}\mathbf{F}$, the modified right Cauchy-Green tensor $\overline{\mathbf{C}} = J^{-2/3}\mathbf{C}$ and the volume ratio $J = \det\mathbf{F}$. Here $\mathbf{I}$ denotes the identity tensor and the material gradient operator

[3] For simplicity a single discontinuity is assumed to demonstrate the underlying kinematics. However, according to the numerical schema described below, we can handle multiple non-interacting discontinuities.

is defined by $\mathrm{Grad}\,(\bullet) = \partial(\bullet)/\partial\mathbf{X}$. In addition, δ_d denotes the *Dirac-delta* functional with $\delta_\mathrm{d} = 0$ and $\delta_\mathrm{d} = \infty$ for $\mathbf{X} \notin \partial\Omega_{0\,\mathrm{d}}$ and $\mathbf{X} \in \partial\Omega_{0\,\mathrm{d}}$, respectively.

In order to provide a discrete constitutive description of the cohesive traction acting on the displacement discontinuity, we introduce the gap displacement $\widehat{\mathbf{u}}(\mathbf{X})$ and the average displacement $\overline{\mathbf{u}}(\mathbf{X})$ connected to a point $\overline{\mathbf{x}}$ on the fictitious spatial discontinuity $\partial\Omega_\mathrm{d}$, see Fig. 1. The spatial orientation of $\partial\Omega_\mathrm{d}$ in $\overline{\mathbf{x}}$ is defined by the unit normal vector $\overline{\mathbf{n}}$, i.e. the weighted push-forward of $\mathbf{N}$. The fictitious spatial discontinuity $\partial\Omega_\mathrm{d}$ is assumed to be in the middle between the two physical surfaces $\partial\Omega_{\mathrm{d}+}$ and $\partial\Omega_{\mathrm{d}-}$.

According to the introduced kinematics, a referential point $\mathbf{X}$ maps into the two spatial points $\mathbf{x}_+$ and $\mathbf{x}_-$, which uniquely defines the gap and average displacements of the discontinuity, i.e. $\widehat{\mathbf{u}} = \mathbf{x}_+ - \mathbf{x}_-$ and $\overline{\mathbf{u}} = (\mathbf{x}_+ + \mathbf{x}_-)/2$, respectively (see Fig. 1).

Variational formulation

The variational formulation for a quasi-static finite element model is based on a single-field variational principle [17], [12], i.e.

$$\int_{\Omega_0} \mathrm{Grad}\,\delta\mathbf{u} : \mathbf{P}(\mathbf{F})\mathrm{d}V - \delta\Pi^{\mathrm{ext}}(\delta\mathbf{u}) = 0, \tag{2}$$

where $\mathbf{P}(\mathbf{F})$ and $\delta\mathbf{u}$ denote the first Piola-Kirchhoff stress tensor and the admissible variation of the displacement field $\mathbf{u}$, respectively. According to the introduced displacement field its admissible variation reads $\delta\mathbf{u} = \delta\mathbf{u}_\mathrm{c} + \mathcal{H}\delta\mathbf{u}_\mathrm{e}$, and after some algebraic manipulations, we get the two spatial variational statements

$$\left.\begin{aligned} \int_{\Omega_-} \mathrm{sym}(\mathrm{grad}_\mathrm{c}\delta\mathbf{u}_\mathrm{c}) : \boldsymbol{\sigma}_\mathrm{c}\mathrm{d}v + \int_{\Omega_+} \mathrm{sym}(\mathrm{grad}_\mathrm{e}\delta\mathbf{u}_\mathrm{c}) : \boldsymbol{\sigma}_\mathrm{e}\mathrm{d}v - \delta\Pi^{\mathrm{ext}}_\mathrm{c}(\delta\mathbf{u}_\mathrm{c}) = 0, \\ \int_{\Omega_+} \mathrm{sym}(\mathrm{grad}_\mathrm{e}\delta\mathbf{u}_\mathrm{e}) : \boldsymbol{\sigma}_\mathrm{e}\mathrm{d}v + \int_{\partial\Omega_\mathrm{d}} \mathbf{t}\cdot\delta\mathbf{u}_\mathrm{e}\mathrm{d}s - \delta\Pi^{\mathrm{ext}}_\mathrm{e}(\delta\mathbf{u}_\mathrm{e}) = 0, \end{aligned}\right\} \tag{3}$$

where $\mathrm{d}v$ and $\mathrm{d}s$ are the spatial volume and surface elements, and $\delta\Pi^{\mathrm{ext}}_\mathrm{c}$ and $\delta\Pi^{\mathrm{ext}}_\mathrm{e}$ are external contributions (virtual external potential energy) which refer to the domains Ω_{0-} and Ω_{0+}, respectively. Here, $\boldsymbol{\sigma}_\mathrm{c} = J_\mathrm{c}^{-1}\mathbf{P}(\mathbf{F}_\mathrm{c})\mathbf{F}_\mathrm{c}^\mathrm{T}$ and $\boldsymbol{\sigma}_\mathrm{e} = J_\mathrm{e}^{-1}\mathbf{P}(\mathbf{F}_\mathrm{e})\mathbf{F}_\mathrm{e}^\mathrm{T}$ denote the Cauchy stress tensors and $\mathbf{t} = \mathbf{T}\mathrm{d}S/\mathrm{d}s$ is the Cauchy traction vector associated with a fictitious discontinuity $\partial\Omega_\mathrm{d}$, and $\mathbf{T}$ is the first Piola-Kirchhoff traction vector associated with the related fictitious discontinuity $\partial\Omega_{0\,\mathrm{d}}$. The spatial gradients in (3) are defined according to $\mathrm{grad}_\mathrm{c}(\bullet) = \mathrm{Grad}\,(\bullet){\mathbf{F}_\mathrm{c}}^{-1}, \mathrm{grad}_\mathrm{e}(\bullet) = \mathrm{Grad}\,(\bullet){\mathbf{F}_\mathrm{e}}^{-1}$ and $\mathrm{sym}(\bullet) = ((\bullet) + (\bullet)^\mathrm{T})/2$ furnishes the symmetric part of $(\bullet)$. For a consistent linearization of the statements (3), we refer to [8] and references therein.

Constitutive formulations

In order to apply the above introduced variational statements, we have to provide continuous and discontinuous constitutive formulations, i.e. the descriptions of the bone tissue as a bulk material, and the fracture. Note that we assume both constitutive formulations to be independent from each other, although some relationship can be provided [19].

Bulk constitutive model. Bone tissue is anisotropic, inhomogeneous and shows time-dependent effects [4]. However, in the present work, we assume isotropic and hyperelastic mechanical properties of bone tissue, and use a simple *neo-Hookean* model to capture its main mechanical behavior. In particular, the strain-energy function $\Psi = \kappa(\ln J)^2/2 + \mu(\mathbf{I} : \overline{\mathbf{C}} - 3)/2$ is proposed, where κ and μ characterize the referential bulk and shear moduli of the bone, respectively.

Cohesive constitutive model. The complex irreversible changes in the microstructure due to bone fracture are lumped into a (discrete) cohesive zone, which is mechanically defined by the cohesive potential ψ. The justification for the application of the theory of cohesive zones rather than assuming sharp crack tips, is motivated by pronounced nonlinear zones, as recently discussed in [29], [30].

Within this work we employ an isotropic cohesive model of the type [10]

$$\psi(\widehat{\mathbf{u}} \cdot \widehat{\mathbf{u}}, \delta) = \frac{T_0}{2\delta} \exp(-a\delta) i_1, \tag{4}$$

where $i_1 = \widehat{\mathbf{u}} \cdot \widehat{\mathbf{u}}$ is the first invariant of the symmetric tensor $\widehat{\mathbf{u}} \otimes \widehat{\mathbf{u}}$. Here T_0 denotes the cohesive strength of bone tissue, and the non-negative parameters a and δ characterize the softening properties of the cohesive zone. In particular, $\delta \in [0, \infty[$ denotes an internal damage variable recording the history of the failure process. According to the procedure by Coleman and Noll [3], the first Piola-Kirchhoff traction is defined by

$$\mathbf{T} = \frac{\partial \psi}{\partial \widehat{\mathbf{u}}} = \frac{T_0}{\delta} \exp(-a\delta) \widehat{\mathbf{u}}. \tag{5}$$

Moreover, the introduction of the mode I fracture energy $\mathcal{G}_I = \int_0^\infty \mathbf{T} \cdot \mathrm{d}\widehat{\mathbf{u}} = \int_0^\infty T_0 \delta^{-1} \exp(-a\delta) \widehat{u} \mathrm{d}\widehat{u} = T_0/a$ of model (4) gives the relation

$$a = \frac{T_0}{\mathcal{G}_I} \tag{6}$$

for the introduced material parameter a.

In order to complete the proposed cohesive description, a damage surface $\phi(\widehat{\mathbf{u}}, \delta) = |\widehat{\mathbf{u}}| - \delta = 0$ in the 3-D gap displacement space is introduced, and it is assumed that $\dot{\delta} = |\dot{\widehat{\mathbf{u}}}|$ captures the evolution of the internal (damage) variable δ.

The proposed model has the advantage that the initial elastic stiffness is infinite. Therefore, the model falls within the regime of 'initially rigid' damage models in which an interface is inactive until the traction across it reaches a critical level, basically the the cohesive strength T_0. However, by implementing the model, one has to circumvent the unbounded initial stiffness, and hence a 'quasi-initially rigid' damage model is implemented [8].

Finally it needs to be emphasized that the proposed cohesive formulation includes only objective quantities, and hence it is objective itself. In particular, $i_1 = \widehat{\mathbf{u}}\cdot\widehat{\mathbf{u}} = \mathbf{Q}\widehat{\mathbf{u}}\cdot\mathbf{Q}\widehat{\mathbf{u}} = i_1^+$ and $\dot{\delta} = \dot{\widehat{\mathbf{u}}}\cdot\dot{\widehat{\mathbf{u}}} = \mathbf{Q}\dot{\widehat{\mathbf{u}}}\cdot\mathbf{Q}\dot{\widehat{\mathbf{u}}} = \dot{\delta}^+$ hold, where $\mathbf{Q}$ denotes a proper orthogonal tensor superimposed on the current configuration, and $(\bullet)^+$ denotes the related quantity $(\bullet)$ in the rotated frame.

2.2 Finite element implementation

In this section we sketch the numerical implementation of the introduced continuum mechanical framework, which allows a mesh-independent propagation of bone fracture without any re-meshing of the considered domain. In particular, the underlying PUFEM, the crack initialization criterion and the crack tracking algorithm are briefly discussed.

Partition of unity finite element method (PUFEM)

According to PUFEM, the displacement field $\underline{\mathbf{u}}$ can be interpolated as[4]

$$\underline{\mathbf{u}} = \sum_{i=1}^{n_{\text{elem}}} N^I \underline{\mathbf{u}}_{I\,\text{c}} + \mathcal{H} \sum_{i=1}^{n_{\text{elem}}} N^I \underline{\mathbf{u}}_{I\,\text{e}}, \tag{7}$$

where N^I are the standard (polynomial) finite element shape functions and n_{elem} denotes the number of nodes of the considered finite element. In (7), regular and enhanced nodal displacements are denoted by $\mathbf{u}_{I\,\text{c}}$ and $\mathbf{u}_{I\,\text{e}}$.

Interpolation (7) in conjunction with the variational statements (3) render the following linearized algebraic set of equations for a particular finite element node I, i.e.

$$\begin{bmatrix} \underline{\mathbf{K}}_{\mathbf{u}_\text{c}\,\mathbf{u}_\text{c}} & \underline{\mathbf{K}}_{\mathbf{u}_\text{c}\,\mathbf{u}_\text{e}} \\ \underline{\mathbf{K}}_{\mathbf{u}_\text{e}\,\mathbf{u}_\text{c}} & \underline{\mathbf{K}}_{\mathbf{u}_\text{e}\,\mathbf{u}_\text{e}} \end{bmatrix}^I_{i-1} \begin{bmatrix} \Delta\underline{\mathbf{u}}_\text{c} \\ \Delta\underline{\mathbf{u}}_\text{e} \end{bmatrix}^I_i = \begin{bmatrix} \underline{\mathbf{f}}^{\text{ext}}_{\mathbf{u}_\text{c}} \\ \underline{\mathbf{f}}^{\text{ext}}_{\mathbf{u}_\text{e}} \end{bmatrix}^I_{i-1} - \begin{bmatrix} \underline{\mathbf{f}}^{\text{int}}_{\mathbf{u}_\text{c}} \\ \underline{\mathbf{f}}^{\text{int}}_{\mathbf{u}_\text{e}} \end{bmatrix}^I_{i-1}, \tag{8}$$

where i, $i-1$ denote the iteration steps associated with a global Newton iteration, and $\Delta\underline{\mathbf{u}}_\text{c}$ and $\Delta\underline{\mathbf{u}}_\text{e}$ denote the increments of the regular and the enhanced nodal displacements, respectively.

In (8) the vectors $\underline{\mathbf{f}}^{\text{ext}}_{\mathbf{u}_\text{c}}$, $\underline{\mathbf{f}}^{\text{ext}}_{\mathbf{u}_\text{e}}$ and $\underline{\mathbf{f}}^{\text{int}}_{\mathbf{u}_\text{c}}$, $\underline{\mathbf{f}}^{\text{int}}_{\mathbf{u}_\text{e}}$ denote external and internal nodal force vectors, respectively. For a detailed derivation of the contributions

[4] Characters indicated by underlines denote the matrix notation of the associated tensor or vector. For example, $\underline{\mathbf{u}}$ is the matrix representation of vector $\mathbf{u}$.

$\underline{\mathbf{K}}_{(\bullet\bullet)}$ (with the abbreviation $(\bullet\bullet)$ for $\mathbf{u}_c\mathbf{u}_c$, $\mathbf{u}_c\mathbf{u}_e$, $\mathbf{u}_e\mathbf{u}_c$, $\mathbf{u}_e\mathbf{u}_e$) to the nodal stiffness tensor, as introduced in (8), see [10] and references therein. Finally, we emphasize that the cohesive traction $\mathbf{T}$ causes an off-diagonal contribution $\underline{\mathbf{K}}_{\mathbf{u}_e\mathbf{u}_c}$ and a diagonal contribution $\underline{\mathbf{K}}_{\mathbf{u}_e\mathbf{u}_e}$ to the element stiffness matrix, see [7], [8]. The proposed PUFEM with linear and quadratic tetrahedral elements have been implemented into the multi-purpose finite element analysis program FEAP, [25].

Finally, it is important to note that PUFEM has a proper variational basis, and it is able to accurately capture strong discontinuity kinematics, which strengthen the numerical reliability. This is not the case for the class of finite element implementations based on the **E**nhanced **A**ssumed **S**train (EAS) method, which can either capture strong discontinuity kinematics or be derived from a variational formulation but none of the EAS-formulations can combine both requirements, see the comparative study in [7], and references therein. This might be the reason that even the non-symmetric version of the EAS-formulation shows ambiguous results for finite strain 3-D problems [7].

Crack initialization criterion

Bone can be regarded as a quasi-brittle material, and hence we assume that the fracture is driven by a (non-local) Rankine criterion. The Rankine criterion in a material point is based on an averaged stress, which is computed over all elements with their center in a sphere with a characteristic radius $\overline{R}$ [9]; a user-specified parameter. It is known that non-local failure criteria can lead to non-physical crack initialization, in particular at stress (strain) singularities [24], however, based on our numerical experience, this limitation turns out to be of minor importance from a practical point of view. Note that the crack initialization criterion defines only the geometry of the discontinuity, while the mechanics of the failure process is characterized by the underlying cohesive law, which is non-local anyway.

Finally we would like to emphasize that the (non-local) Rankine criterion is the simplest approach, and more complex and anisotropic failure criteria for bone tissue have been proposed in the literature as well, see, for example, the Tsai-Wu criterion, as discussed in [4] and references therein.

Crack tracking algorithm

A crucial task in taking advantage of PUFEM is the geometrical representation of the crack surface and the tracking of its evolution. Especially for 3-D problems, the development of crack tracking algorithms is an active research area in computational mechanics, and *local* [16] and *global* [20] crack tracking schemes have been recently proposed in the literature.

In contrast to those propositions, we use here the recently proposed (partially local) two-step predictor-corrector algorithm for tracking 3-D crack paths [8], which is summarized in the following. We assume now that the

failure criterion of the k-th finite element (located at the crack tip) is met. Hence, the related discontinuity needs to be appended to the existing crack. The orientation of the discontinuity is defined by the failure criterion, and the existing crack defines the location in space so that from a geometrical point of view the problem is of Runge-Kutta type. However, the straightforward application of the non-local Rankine criterion in a 3-D geometric setting may cause the development of a geometrically incompatible ('bumpy') crack surface, which may lead to (non-physical) crack bifurcations. In order to circumvent these topological difficulties, we have recently proposed a two-step predictor-corrector algorithm, where the predictor step defines discontinuities according to the (non-local) Rankine criterion, and the corrector step is based on a surface smoothing strategy [9]. In particular, the corrector step draws in non-local information of the existing discontinuities in order to predict a 'smooth'[5] 3-D crack surface. This is realized by fitting a polynomial surface $Z(X,Y)$ locally to the predicted crack surface, i.e. the crack surface defined by the predictor step. To this end $n_{R^\star}$ points on the crack surface (located in the vicinity $\Omega_0^\star$ of the k-th finite element) are considered, and the coefficients of the polynomial surface are defined by minimizing the least-square problem

$$\Phi = \sum_{i=1}^{n_{R^\star}} (Z_i - Z(X_i, Y_i))^2 \rightarrow \text{MIN}, \tag{9}$$

where X_i, Y_i, Z_i denote the components of the considered points on the crack surface. In the present analysis $\Omega_0^\star$ is assumed to be a sphere of radius $R^\star$.

Subsequently, according to [9], the orientation of the discontinuity $\mathbf{N}$ in the k-th finite element is adapted (corrected) to the normal onto the polynomial surface $Z(X,Y)$. This leads to a 'smooth' prediction of the crack surface and avoids topological problems, which are inherent in the predictor step.

The proposed two-step tracking algorithm has been implemented in a separate user macro and linked to FEAP [25]. The macro is executed after each mechanical loading step, and the user has to specify $R^\star$ and the degree of the polynomial surface; the current implementation supports linear and quadratic surfaces.

3 Numerical example

The numerical example described in this section follows closely the experimental investigation documented in [21]. Therein, the proximal part of a femur with an approximate length of 20 cm was subjected to loading with 0.5 mm/s until failure occurred. The distal part of the femoral bone was fixed in a device, and a compressive load was applied on the femoral head.

[5] Note that the proposed geometrical representation of the crack leads to a $\mathcal{C}^1$ continuous surface, and the word 'smooth' has no mathematical meaning here.

For the present numerical study, we used the geometry of the proximal part of a 'standardized femur' [26]. The geometry was discretized using tetrahedral finite elements and the mesh was generated with the software-package NETGEN, [23]. The finite element model consists of 5212 nodes. The mesh was slightly refined at those regions where failure was expected; known from experimental studies.

We assume a homogeneous elastic stiffness, defined by the referential Young's modulus of 10.0 GPa and the referential Poisson's ratio of 0.35, which falls within the range of reported mechanical properties of femoral bone [1], [4]. These values correspond to referential bulk and shear moduli of $\kappa = 11.1$ GPa and $\mu = 3.7$ GPa, respectively. In order to account for the different cohesive strength T_0 in the bone we assume a value of 7.0 MPa for the epiphysis and metaphysis and 70.0 MPa for the diaphysis. In addition, we assume that T_0 varies linearly between these two values within a 3.0 cm transition zone. According to the values provided in [27], [29] the mode I fracture energy was approximated to be a constant with value $\mathcal{G}_I = 2.0$ kJ/m^2. With this value the cohesive parameter a may be computed. The parameter a varies from 3.5 mm^{-1} in the epiphysis to 35.0 mm^{-1} in the diaphysis.

In order to use the proposed crack initialization criterion, the introduced parameter $\overline{R}$ was set to $3.0\sqrt[3]{V_\mathrm{e}}$, where V_e denotes the referential volume of the finite element under consideration. In addition, the crack tracking algorithm was used by a linear polynomial surface and $R^\star$ was taken to be $3.0\sqrt[3]{V_\mathrm{e}}$ in the present analysis.

In the finite element model all nodes of the cross-section at the bottom, where the shaft is mounted to the testing machine [21], were fixed. At the very top of the femoral head a dead load was applied at 9 nodal points. The magnitude of the load at the different nodal points is the same and the related load direction passes through the center of fixed cross-section at the bottom.

A displacement controlled arc-length method was applied to solve the nonlinear problem [5], where the axial displacement increment of the very top point of the femoral head controls the load-level parameter of the continuation method. In the present computation we considered 215 load steps, which took about 4 hours CPU on a PC with 1.8 GHz, 1.0 Gb RAM and one Pentium V processor.

The computed load-displacement curve is provided in Fig. 2, where, in particular, the magnitude of the applied compressive load is plotted against the axial displacement which occurs at the very top nodal point of the femoral head. At point (a) bone failure starts do develop, however, the structure remains stable until point (b), where the limit load of 8.88 kN is reached. Exceeding this point a sudden drop of the load is observed until the structure recovers at point (c) where it can withstand slightly increasing load again. It needs to be emphasized that the predicted limit load of 8.88 kN coincides nicely with the experimental value of 8.4 kN, documented in [21]. However, this agreement should be seen with caution, because neither the specific geom-

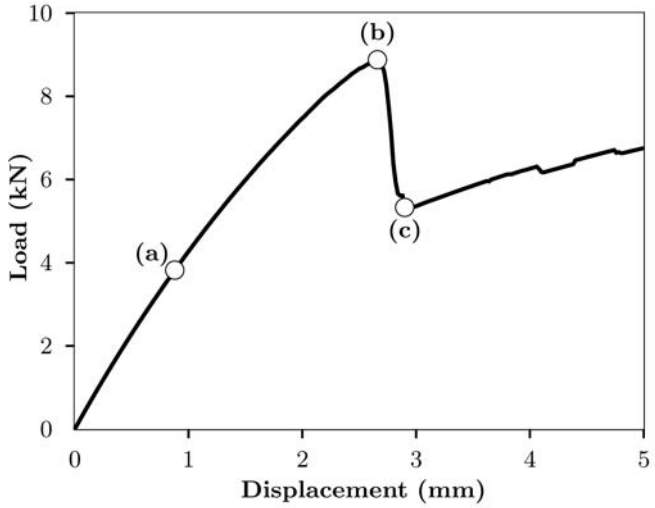


Fig. 2. Predicted load-displacement response of the proximal part of the femur under compressive load: states (a), (b) and (c) denote onset of failure, limit load and recovery of the structure, respectively.

etry of the femur nor the information about the morphology of the bone were available.

The computed maximum principal Cauchy stress distribution at the three states (a)–(c) are illustrated in Fig. 3. Apart from the stress concentration at the crack tip, high tensile stresses appear due to bending of the shaft of the femur. However, the high strength of the cortical bone prevents crack formation in that region.

The predicted crack formation is summarized in Fig. 4, where Fig. 4(a) shows the fully developed slightly curved 3-D crack surface, and Fig. 4(b) illustrates the good agreement with the experimental results taken from [21].

4 Summary and Conclusion

Fracture of the proximal part of the femur is frequently observed in the elderly population, and a better understanding of this type of bone failure can help to improve preventive treatments. In particular, numerical fracture simulations can provide further insight into the onset and progress of femoral fracture under varying loading conditions. Bone fracture is known to develop significant nonlinear zones ahead the crack tip and cannot be described by linear fracture mechanics [29], [30].

In view of its proper variational basis and the capability to capture strong discontinuity kinematics, we applied PUFEM and combined the method with the cohesive crack concept [28], [8]. An important task to be processed when

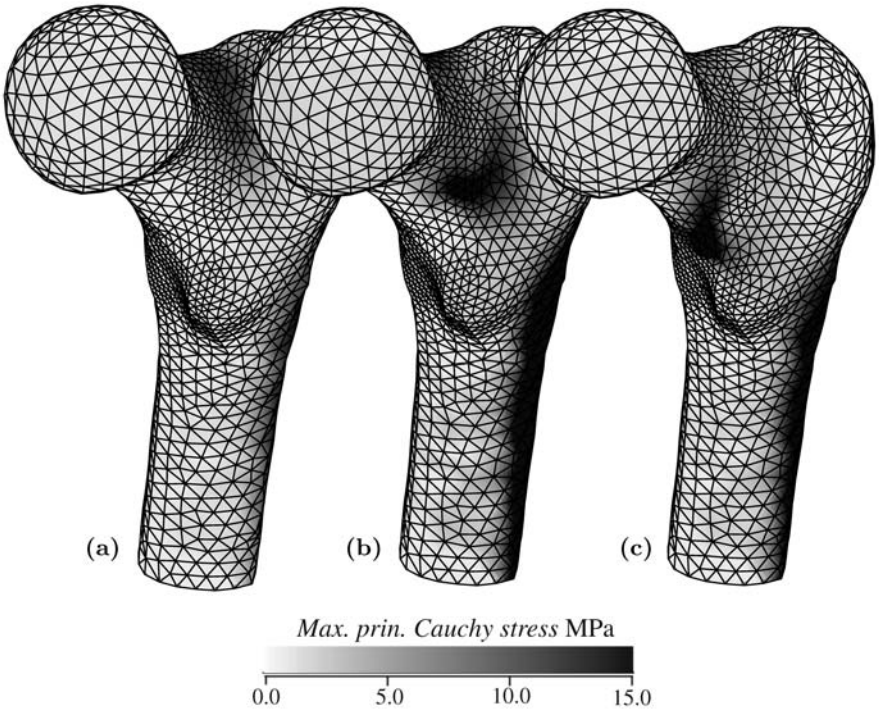


Fig. 3. Distribution of the maximum principal Cauchy stresses at states (a), (b) (c), which denote onset of failure, limit load and recovery of the structure, respectively (compare also with Fig. 2).

using this combined approach, in particular within a 3-D geometric setting, is the implementation of an appropriate crack tracking algorithm. We used the (partially local) two-step predictor-corrector algorithm, as proposed in [9].

The combination of PUFEM with the cohesive crack concept and the algorithm developed for tracking 3-D crack paths turns out to be a very efficient and powerful approach to investigate bone fracture. In the present work we studied the failure mechanism of the proximal part of a 'standardized femur' under compressive load, as experimentally studied in [21]. We used a 3-D finite element model which was generated by means of tetrahedral elements and slightly refined at those regions where failure was expected. For the homogeneous bulk material we assumed linear elasticity, while the fracture properties were assumed to vary between the epiphysis and the diaphysis. We used a constant value of the mode I fracture energy, from which the cohesive parameter a was computed using eq. (6). A displacement controlled arc-length continuation method was used to predict the crack pattern and limit load, which coincide very well with the experimental observations documented in [21].

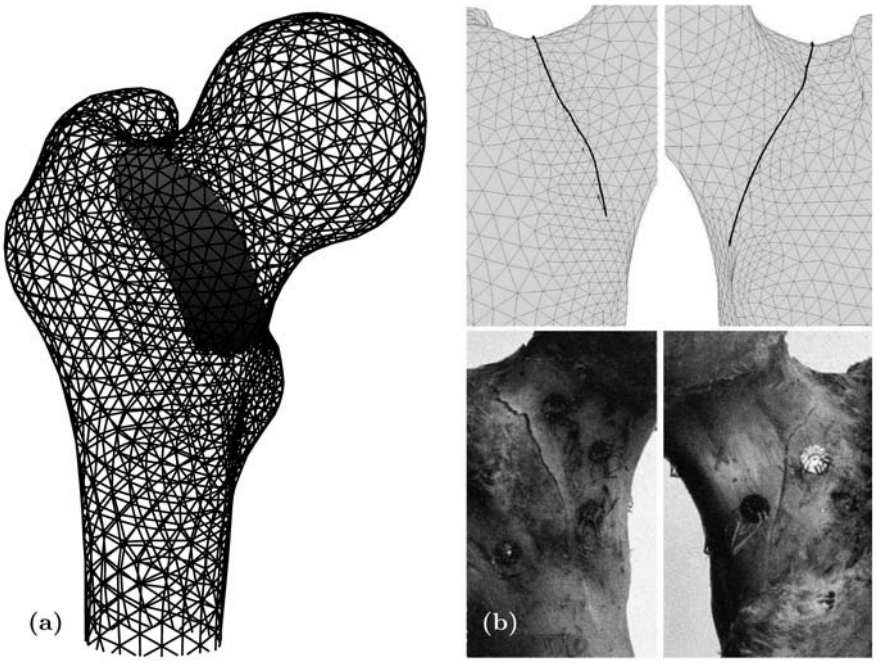


Fig. 4. Predicted crack formation: (a) fully developed slightly curved 3-D crack surface; (b) crack paths on the bone surface and comparison with experimental results taken from [21].

The presented constitutive models of the proximal part of the femur do not account for the complex regional variations of the mechanical properties, as it is indicated by the related changes in bone density. In addition, the application of an isotropic bulk material model, and the assumptions made regarding the crack initialization criterion and the cohesive zone model might limit the predictability of the proposed framework. However, note that the presented framework is general enough to allow for several refinements of the constitutive formulations.

References

1. Abè H, Hayashi K, Sato M, editors, (1996) Data Book on Mechanical Properties of Living Cells, Tissues, and Organs. Springer
2. Armero F, Garikipati K (1996) Int. J. Solids Structures 33:2863–2885
3. Coleman BD, Noll W (1963) Arch. Rat. Mech. Anal. 13:167–178
4. Cowin SC, editor, (2001) Bone mechanics handbook. CRC Press
5. Crisfield MA (1991) Non-linear Finite Element Analysis of Solids and Structures, Essentials, Volume 1. John Wiley & Sons

6. Ford CM, Keaveny TM, Hayes WC (1996) J. Bone Miner. Res. 11:377–383
7. Gasser TC, Holzapfel GA (2003) Comput. Meth. Appl. Mech. Eng. 192:5059–5098
8. Gasser TC, Holzapfel GA (2005) Comput. Meth. Appl. Mech. Eng. 194:2859–2896
9. Gasser TC, Holzapfel GA (2006) Comput. Meth. Appl. Mech. Eng. 195:5198–5219
10. Gasser TC, Holzapfel GA (2006) Eur. J. Mech. A/Solids 25:617–633
11. Gómez-Benito MJ, García-Aznar JM, Kuiper JH, Doblaré M (2005) J. Theor. Biol. 235:105–119
12. Holzapfel GA (2000) Nonlinear Solid Mechanics. A Continuum Approach for Engineering. John Wiley & Sons
13. Keyak JH, Rossi SA (2000) J. Biomech. 33:209–214
14. Lotz JC, Cheal EJ, Hayes WC (1991) J. Biomech. Eng. 113:353–360
15. Lotz JC, Cheal EJ, Hayes WC (1991) J. Biomech. Eng. 113:361–365
16. Moës N, Gravouil A, Belytschko T (2002) Int. J. Numer. Meth. Eng. 53:2549–2568
17. Ogden RW (1997) Non-linear Elastic Deformations. Dover
18. Oliver J (1996) Int. J. Numer. Meth. Eng. 39:3575–3600
19. Oliver J (2000) Int. J. Solids Structures 37:7207–7229
20. Oliver J, Huespe AE, Samaniego E, Chaves EWV (2004) Int. J. Numer. Anal. Meth. Geomech. 28:609–632
21. Ota T, Yamamoto I, Morita R (1999) J. Bone Miner. Metab. 17:108–112
22. Pietruszczak S, Inglis D, Pande GN (1999) J. Biomech. 32:1071–1079
23. Schöberl J (2002) NETGEN – A 3D tetrahedral mesh generator – Version 4.2. University Linz, Austria
24. Simone A, Askes H, Sluys LJ (2004) Int. J. Solids Structures 41:351–363
25. Taylor RL (2002) FEAP – A Finite Element Analysis Program, Version 7.4 User Manual. University of California at Berkeley, Berkeley, California
26. Viceconti M, Casali M, Massari B, Cristofolini L, Bassini S, Toni A (1996) J. Biomech. 29:1241
27. Wang X, Puram S (2004) Ann. Biomed. Eng. 32:123–135
28. Wells GN, Sluys LJ (2001) Int. J. Solids Structures 50:2667–2682
29. Yang QD, Cox BN, Nalla RK, Ritchie RO (2006) Biomaterials 27:2095–2113
30. Yang QD, Cox BN, Nalla RK, Ritchie RO (2006) Bone 38:878–887

Application of X-FEM to 3D Real Cracks and Elastic-Plastic Fatigue Crack Growth

A. Gravouil[1], A. Combescure[1], T. Elguedj[1], E. Ferrié[2], J.-Y. Buffière[2] and W. Ludwig[2]

[1] LaMCoS UMR CNRS 5514 INSA Lyon, 20 allée des sciences, 69621 Villeurbanne Cedex, France; anthony.gravouil@insa-lyon.fr
[2] MATEIS UMR CNRS 5510 INSA Lyon, 20 allée des sciences, 69621 Villeurbanne Cedex, France

Summary. In a general point of view, X-FEM plus level set representation of the interfaces reveals to be of great interest in the aim to couple experimental data with numerical simulations. This can be highly illustrated in the case of 3D fatigue crack growth simulations where the initial 3D "real crack" is extracted from tomographic images. The experimentally observed fatigue crack propagation is compared to numerical simulations. Good agreement is found when a linear variation of closure stress along the crack front is taken into account in the "3D crack propagation law" used for the simulation. Furthermore, in order to take into account plasticity during fatigue crack growth, one develops an augmented Lagrangian formulation in the X-FEM framework that is able to deal with elastic-plastic crack growth with treatment of contact and friction. Numerical issues such as contact treatment and numerical integration are addressed, and finally numerical examples are shown to validate the method.

Key words: X-FEM, level sets, X-ray microtomography, fatigue crack growth, plasticity, contact.

1 Introduction

It is now well established that the eXtended Finite Element Method is of great interest for evolving discontinuities, in particular for industrial applications [13, 14]. Indeed, no initial mesh and remeshing techniques are needed during the evolution of the interfaces [2, 3, 12]. Furthermore, in a more general point of view, X-FEM plus level set representation of the interfaces reveals to be of great interest in the aim to couple experimental data with numerical simulations [1, 2, 12]. This can be highly illustrated in the case of 3D fatigue crack growth simulations where the initial 3D "real crack" is extracted from tomographic images with a spatial resolution of the order of 1 µm [4]. In such a case, X-FEM is coupled with a level set representation of the crack and a

Alain Combescure et al. (eds.), IUTAM Symposium on Discretization Methods for Evolving Discontinuities, 213–231.

specific algorithm is developed in the aim to define the initial level sets from tomographic images. One can also notice that it is not necessary to impose the compatibility between the meshes of the structure and the support of the level sets. Furthermore, a robust and accurate technique is proposed to compute the stress intensity factors [1]. Based on experimental observations and quantitative analysis of crack propagation in the bulk, a "3D" crack growth law can be established to predict the crack front shape evolution in the bulk of the material. From a more general point of view, this study shows that coupling X-ray microtomography to X-FEM provides a promising tool to assess the 3D behaviour of arbitrary shaped cracks and to perform direct comparisons of "experimental" and "simulated" crack shapes during propagation.

In a second example, it is shown that even with elastic-plastic behaviour coupled with contact and friction, non-remeshing property can be preserved for instance for 2D mixed-mode plastic fatigue crack growth [4, 5]. For that purpose, an augmented Lagrangian formulation in the X-FEM framework for the treatment of contact and friction with elastic-plastic behaviour is proposed. On the one hand, the numerical integration is adapted in order to properly integrate the high order terms in the enrichment basis, and to have a fine knowledge of the stress state around the tip to precisely model plasticity. On the other hand, the strategy of enrichment used for linear elastic X-FEM fatigue simulation has to be modified [5]. In this respect, the integration strategy is designed in order to avoid the projection of stresses and internal variables as the crack evolves to ensure the reliability of the method. Numerical issues are addressed. In particular, one shows the ability of the method to model the phenomenon of crack closure under cyclic tension [6, 9]. Indeed, this can have a great influence on the propagation of the crack with fatigue loading.

2 Application of X-FEM to 3D Real Cracks

2.1 Experimental Observations

The propagation of a semi-elliptical crack in the bulk of an ultrafine-grained Al–Li alloy has been investigated using synchrotron radiation X-ray microtomography. In this material, the studied crack, despite its small dimension, can be considered as a "microstructurally long" and described in the frame of the linear elastic fracture mechanics. The main advantage of using this alloy for the present study is that it exhibits an exceptionally linear crack path compared to ingot metallurgy Al alloys. The ultrafine grains promote homogeneous deformation and prevent crystallographic cracking so that the crack shape is not disturbed by microstuctural features, at least not at the level of the spatial resolution employed in this study. In situ fatigue tests monitored by X-ray microtomography were carried out on beamline ID 191 of the European synchrotron radiation facility (ESRF) in Grenoble, France.

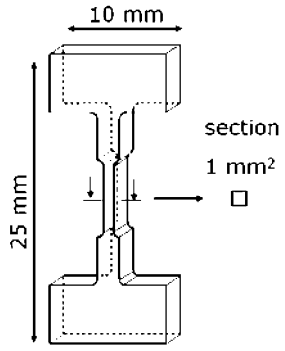


Fig. 1. Geometry of the sample used for the in situ fatigue test monitored by X-ray microtomography

This experimental station is dedicated to high-resolution X-ray imaging and features a highly coherent X-ray beam, a precision mechanics sample stage and a high-resolution detector system. During a 3D tomographic scan, the sample is rotated over 180 in order to acquire 1500 two-dimensional (2D) projection radiographs. From this set of projections, a quantitative 3D map of the attenuation coefficient distribution within the sample is produced by means of a standard filtered backprojection tomographic reconstruction algorithm. The spatial resolution obtained in the reconstructed images is of the order of 1 µm, a value comparable to the resolution of an optical microscope. One can also notice that the 3D scans are done with the maximal load: it ensures that the crack is open. However, even with a locally closed crack, diffraction effect allows to measure a crack opening until 100 nm. The sample was imaged in a dedicated fatigue machine designed at INSA Lyon to perform in situ fatigue test at the ESRF. The mechanical design of the cyclic tension loading mechanism enables operation at cycling frequencies of up to 80 Hz, minimising thereby the cycling time required for a fatigue test.

The geometry of the sample used is represented in Figure 1. A thin (2 µm) rectangular notch, 100 µm wide and 20 µm deep, was produced in the sample using focused ion beam machining. This notch is located at the centre of one of the specimen faces and acts as a crack initiation site.

The in situ fatigue test was performed in air, at constant stress amplitude with $\sigma_{\max} = 220$ MPa, a stress ratio $R = 0.1$, a frequency of 40 Hz and at room temperature. In the fine-grained alloy studied here, the crack size (which is in the range 100–500 µm) is at least 100 times the grain size, and can hence be considered as microstructurally long.

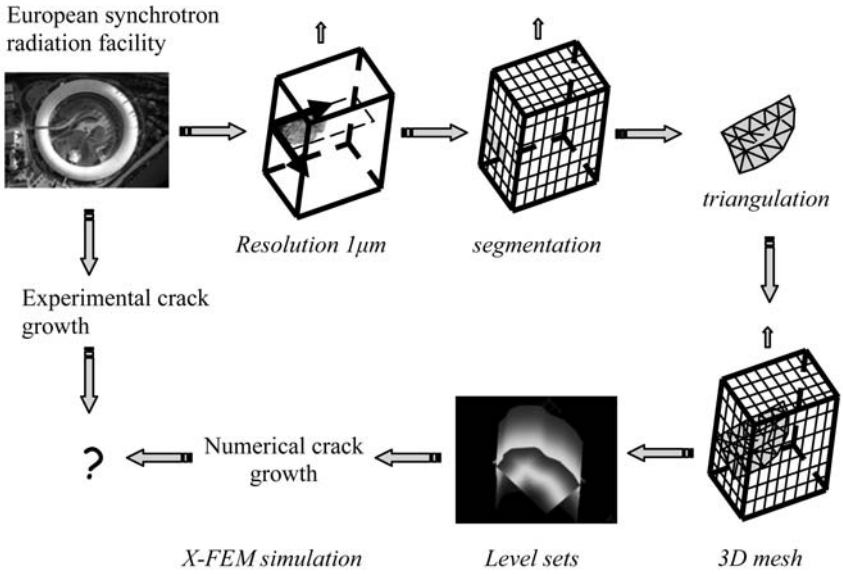


Fig. 2. From X-ray microtomography to X-FEM

2.2 Implicit Representation of the Crack Using the Level Set Method

From the previous experiments, 3D images are obtained and exploited in order to proceed to a simulation of the crack propagation. In this respect, one represents the crack by the use of two level sets.

Indeed, an implicit representation of the crack using the level sets method is very well suited to the eXtended Finite Element Method [1, 2].

The crack front is defined as the intersection between a surface C_{crack} that defines the crack surface and a surface C_{front} that defines the crack front. These two surfaces, defined for a semi-elliptical crack, are displayed in Figure 4. In the general case, these surfaces can have arbitrary 3D shape, and are defined by means of level sets functions. In order to couple the level set representation of the crack with the 3D images which come from X-ray microtomography, a specific strategy is proposed and is summarized in Figures 2 and 3.

The first step consists in a segmentation of the 3D image in order to extract the crack. The result is a set of voxels which defines a volume linked to the crack. The second step consists in the definition of a Boolean field calculated from the segmentation: a voxel outside the crack is initialized to 0 and a voxel inside the crack is initialized to 1. The main difficulty is now to be able to define a surface from this crack volume. In this respect, for the third step, it is needed first to "model" the crack as a surface with a predefined length scale.

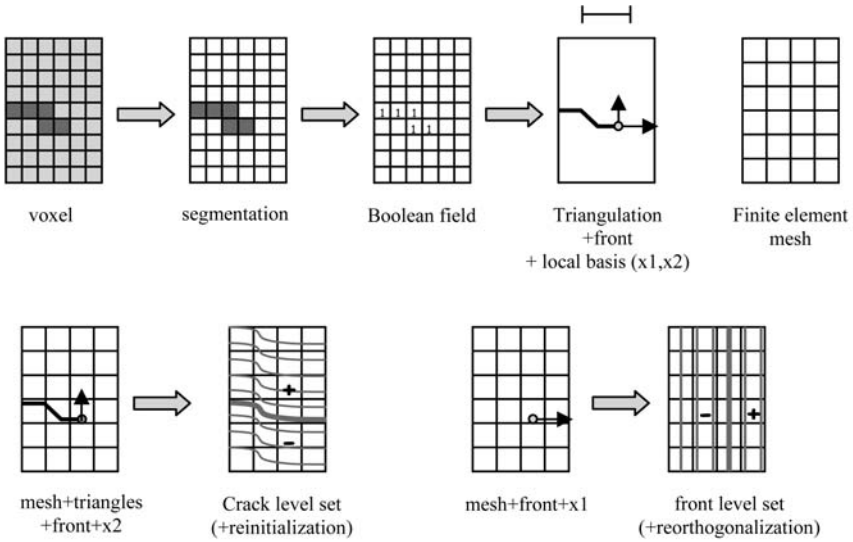


Fig. 3. Initialization of crack and front level sets

As a consequence, a length parameter is defined, and an intermediate triangulation step is done according to this length scale (in practice greater than the thickness of the crack).

Indeed, it seems essential in practice to model the crack as an explicit non-planar surface before to define the level sets. From the set of triangles, the 1D elements of the front are extracted with the local basis along the crack front. The fourth and the fifth steps consist then in defining the crack and the front level sets from the set of triangles and the set of 1D elements with their respective local basis.

In this respect, reinitialization equations (to the sign distance function) and orthogonalization equations proposed in [2] are used according to Figure 3 in order to compute the two initial 3D scalar fields linked to the crack and the front (see Figure 4). In particular, this approach can be done in a narrow band close to the crack.

2.3 Strategy of Enrichment and SIFS Calculation

In FE methods, the presence of cracks in a structure must be taken into account in mesh generation: the mesh must adequately define the crack geometry. Special elements and considerable mesh refinement near the crack tip are necessary to capture accurately the asymptotic displacement fields. The X-FEM alleviates the shortcomings associated with meshing the crack surface by representing the crack geometry using additional functions called enrichment functions. The enrichment method, as described in [1], can be summarised as

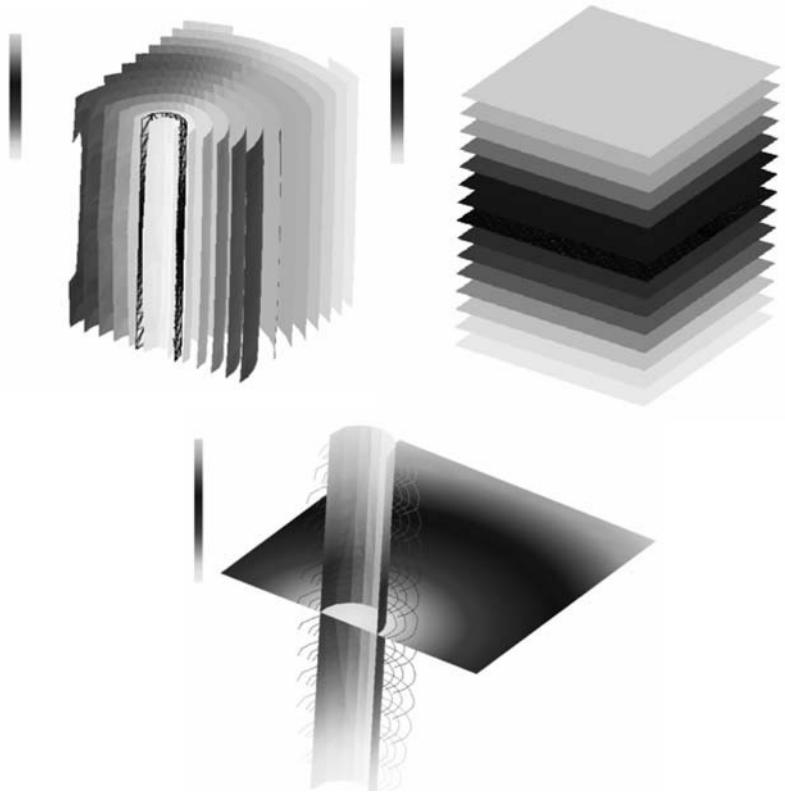


Fig. 4. Representation of the 3D scalar fields linked to the crack and front level sets

follows. The displacement field calculated in the structure "without crack" is locally modified by adding, at specific nodes, the nodal values of the enrichment functions. These functions define the crack geometry by "modelling" the discontinuity introduced by the crack in the displacement field. As the discontinuities at the crack front and along the crack surface are different, two enrichment functions are necessary to model the entire crack [1].

$$H(\mathbf{x}) = \begin{cases} 1 & \text{for } \phi > 0, \\ -1 & \text{for } \phi < 0, \end{cases} \tag{1}$$

$$\{F_J(\mathbf{x})\} \equiv \sqrt{r} \left\{ \sin\frac{\theta}{2}, \cos\frac{\theta}{2}, \sin\frac{\theta}{2}\sin\theta, \cos\frac{\theta}{2}\sin\theta \right\}. \tag{2}$$

The function H, defined as a generalised Heaviside function, models the displacement field along the crack surface. It is used to enrich the nodes of the element that are cut by the crack surface.

The enrichment F_J is used to enrich the nodes of the elements that contain the crack front. F_J consists of a span of functions which incorporate the

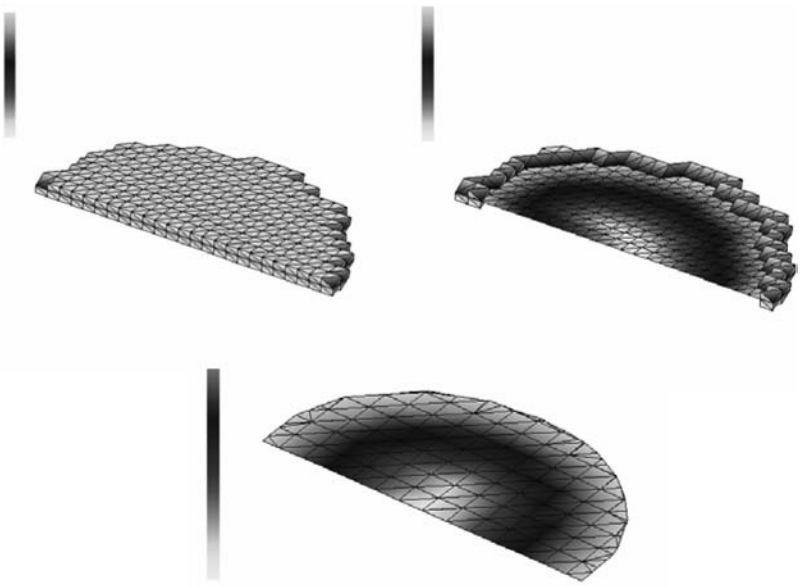


Fig. 5. Set of elements completely cut by the crack; set of elements cut by the front; 2D representation of the crack extracted from the crack level set

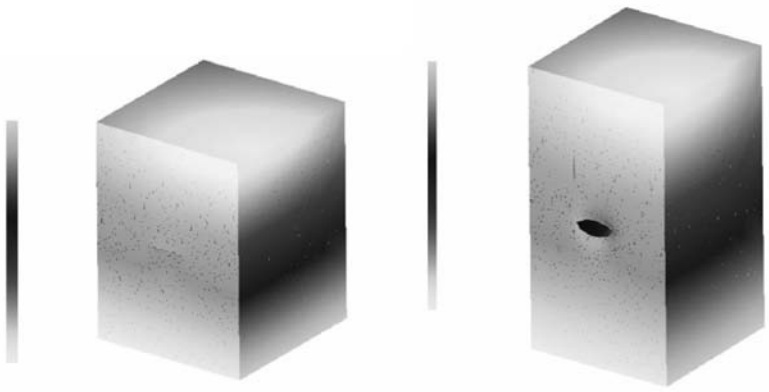


Fig. 6. Representation of the non-deformed and deformed mesh of the specimen submitted to fatigue loading

radial and angular behaviour of the asymptotic crack-tip displacement fields, where the local coordinate system can be extracted from the level sets [2]. Furthermore, the nodal values of the crack level set and the front level set give the precise location of the crack and control whether a node has to be enriched or not (see Figure 5).

Concerning the calculation of the local stress intensity factor values (K_{I}, K_{II} and K_{III}) along the crack front, the interaction integral is used on a J-domain defined in the local basis of the crack front [1, 2]. The J-domain is completely independent of the mesh of the structure and its size depends on the curvature of the crack front and the size of the finite elements cut by the front. On the boundary, due to plane stress effect, the interaction integral is not used and stress intensity factors in the bulk are extrapolated from the bulk to the boundary [4]. In Figure 6, we illustrate the non-deformed mesh and the deformed mesh of the specimen submitted to a fatigue load after 14000 cycles.

3 Development of a 3D Local Paris Law

After 14,000 fatigue cycles, a crack initiated at the notch was detected on a 2D radiograph. The in situ fatigue test was performed at constant stress amplitude with $\sigma_{\max} = 220$ MPa, a stress ratio $R = 0.1$, and a frequency of 40 Hz. A tomographic scan of the part of the specimen containing the crack was acquired and corresponds to the "first scan". During crack propagation, nine scans were recorded, reconstructed and segmented to obtain 3D renditions of the crack at different stages of its evolution. Projections in the $(X–Y)$ plane of seven 3D renditions (chosen among the nine) are displayed in Figure 7. The dotted lines on the 3D crack renditions represent the location of the crack front at the previous step. The position along the crack front is defined by the angle x as shown in Figure 7(a). Qualitative and quantitative information on the evolution of the 3D crack geometry can be obtained.

The crack size at the surface (2c), for $\omega = 0°$, and in the bulk (a), for $\omega = 90°$, are measured on the 3D renditions of the crack. The fatigue crack growth rates at the surface, dc/dN, and in the bulk, da/dN, as a function of the stress intensity factor range $\Delta K = (K_{\max} - K_{\min})$ are shown in Figure 9. The values of $K_{\max}$ and $K_{\min}$ are calculated for $\omega = 0°$ and for $\omega = 90°$ using domain integral techniques [1].

Thus for the same ΔK, the crack propagates faster in the bulk than at the surface implying that, at least for the specimen geometry used here, the crack growth behaviour is anisotropic. Thus, a fatigue crack growth law determined at the surface does not account for the crack behaviour in the bulk. Furthermore, if used for fatigue life calculation, the surface experimental law would lead to a non-conservative prediction.

One possible reason for the observed crack growth anisotropy between the surface and the bulk is variation of the closure stress along the crack front: the crack will then propagate faster in the bulk because the closure stress in this region of the crack front is smaller and hence the effective driving force is higher than at the surface.

Elber [6] proposed a modified Paris equation to account for the effect of closure on the crack growth rate: $da/dN = C(\Delta K_{\mathrm{eff}})^m$ where $\Delta K_{\mathrm{eff}} =$

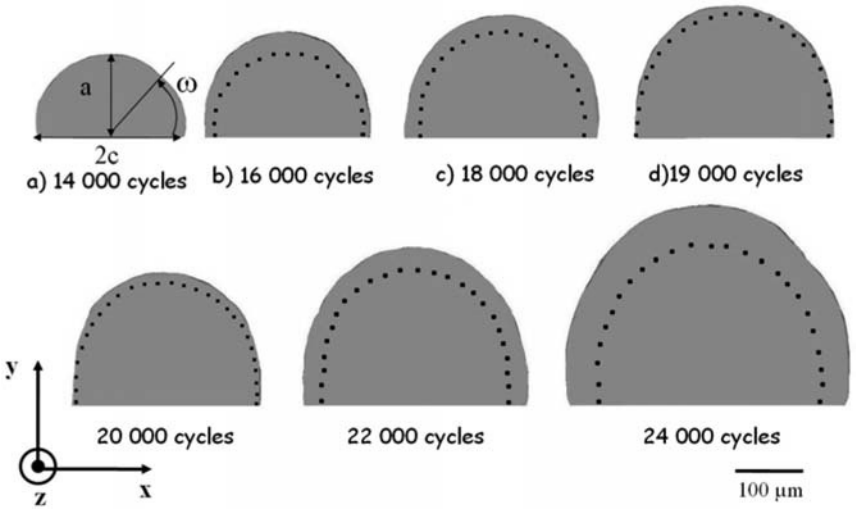


Fig. 7. 3D segmentation operation of the crack from the 3D tomographic image during propagation

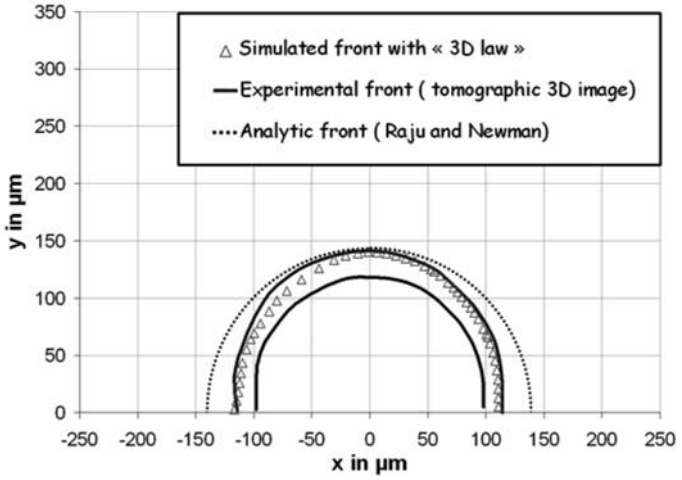


Fig. 8. Comparisons between experimental, numerical and analytical solutions from 14000 to 16000 cycles

$(K_{\max} - K_{\mathrm{op}})$ is the effective stress intensity factor range and K_{op} is the stress intensity factor at which the crack opens. To account for the variation of the closure stress along the crack, we will modify Elber's equation by introducing a new variable ω such that da/dN and ΔK_{eff} vary along the crack front: $da(\omega)/dN = C(K_{\max} - K_{\mathrm{op}}(\omega))^m$.

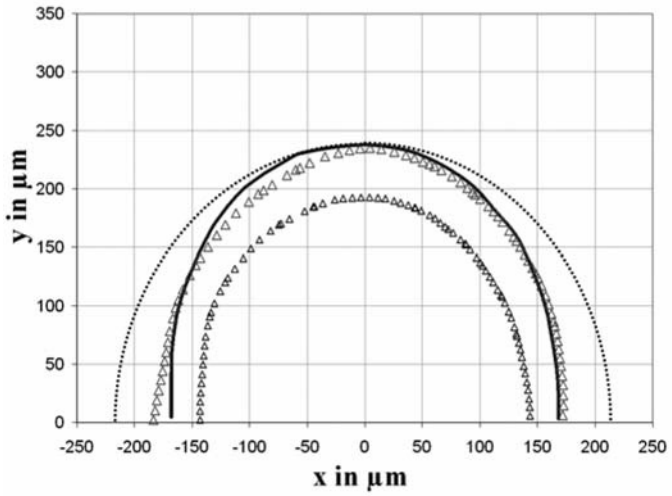


Fig. 9. Comparisons between experimental, numerical and analytical solutions from 20000 to 22000 cycles

As a first approach, the variation of $K_{\mathrm{op}}(\omega)$ is taken to be linear with the angle ω and is therefore given by $K_{\mathrm{op}}(\omega) = a\omega + b$. The value of the constants a and b are determined from the values of K_{op} at the surface $K_{\mathrm{op}}(0°)$ and at the deepest point $K_{\mathrm{op}}(90°)$. Data from the literature are used to determine $K_{\mathrm{op}}(0°)$: the closure response of an Al–Li powder metallurgy alloy, very similar to the alloy studied here, was investigated in [7]. The value of K_{op} measured at the surface of CT specimens, for $R = K_{\max}/K_{\min} = 0.1$, is found to be equal to $0.4K_{\max}$. We assume here that the ratio $K_{\mathrm{op}}(\omega = 0°)/K_{\max}$ remains constant during the fatigue test and that there is no closure at the deepest point for $\omega = 90°$, which gives $K_{\mathrm{op}}(90°) = 0.1K_{\max}(90°)$. The evolution of the closure stress along the crack front is given by the linear relation:

$$K_{\mathrm{op}} = -\frac{0.3K_{\max}}{90°}\omega + 0.4K_{\max}. \tag{3}$$

The 3D crack growth law taking into account the variation of the closure stress along the crack front is given by

$$\frac{da}{dN}(\theta) = C\left(\left(\frac{0.3}{90°}\omega + 0.6\right)K_{\max}\right)^m \tag{4}$$

with $C = 10^{-9.2}$ mm.cycle^{-1}.(MPa.m$^{0.5}$)$^{-m}$ and $m = 3.51$ for the previous aluminium alloy. In Equations (3) and (4), the angle ω, expressed in degrees, is defined for $0° < \omega < 90°$ and $K_{\max}(\omega)$ is obtained from a polynomial interpolation of the values calculated by the X-FEM as described in the previous section. This means that, for the sample geometry investigated here, a single Paris law can be used to predict the observed crack growth anisotropy

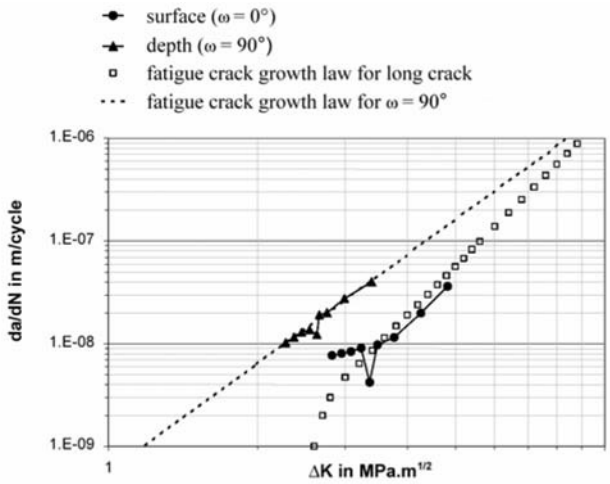


Fig. 10. Three-dimensional effect (difference between surface and bulk)

provided the variation of the closure stress along the crack front is taken into account.

From a more general point of view, this study has shown that coupling X-ray microtomography to X-FEM provides a promising tool to assess the 3D behaviour of arbitrary shaped cracks and to perform direct comparisons of "experimental" and "simulated" crack shapes during propagation. Such data, to the best of the authors' knowledge, are currently lacking in the literature.

In this study, closure effect is taken into account empirically in the Paris law. An extension of this work consists in using explicitly a elastic-plastic behaviour with contact and friction in the simulation. This is the aim of the next section.

4 Extension of X-FEM to Elastic-Plastic Crack Growth

4.1 Strategy of Enrichment

The main purpose of this paragraph is to treat the case of bulk and interface non-linearities in the framework of the eXtended Finite Element Method. The presented method will focus on the case of plasticity combined with frictional contact and is applied to fatigue crack growth analysis. The aim of the method is to use the X-FEM to model the well known fatigue crack closure under cyclic tension phenomenon first observed by Elber [8].

In the case of material non-linearities, several issues have to be addressed. In a previous paper [5] an appropriate elastic-plastic enrichment basis was developed to allow the X-FEM to deal with plasticity. In the present paper this

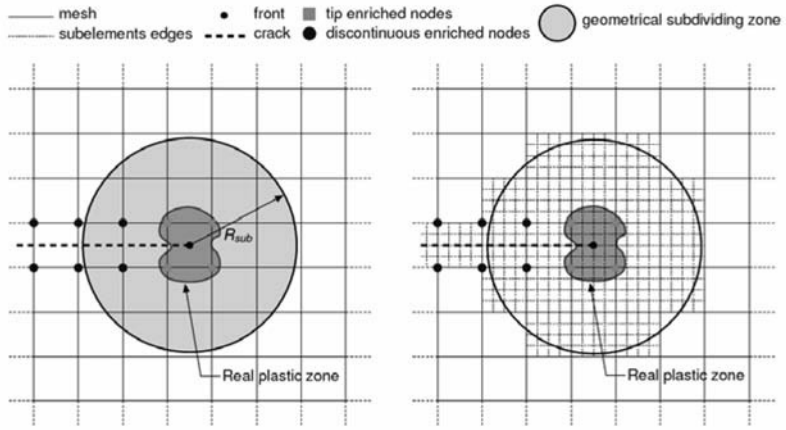


Fig. 11. Geometrical subdividing zone with an estimation of the plastic zone

approach is coupled with the treatment of frictional contact, and the problems associated with plastic crack propagation are explored. The aim of the study is to develop a general strategy for a wide family of elastic-plastic behaviour (kinematics and/or isotropic hardening, etc.). In this approach, neither remeshing nor thermodynamic field interpolation during propagation of the crack are needed. Furthermore, the crack front is enriched with specific basis in order to use sufficiently coarse mesh compared to FEM. In fact, the aim is to be able to take into account closure effect and confined plasticity with a good accuracy during propagation of the crack: indeed, this can have a great influence on the fatigue crack growth. In order to accurately model the confined plasticity, a geometrical subdividing zone is defined with an estimation of the plastic zone which can be smaller than the spatial discretization assumed to be fixed (see Figure 11). In the same way, during the propagation of the crack, one defines new sub-elements between the two configurations according to the estimation of the new plastic zone (see Figure 12): only the Gauss points of the new sub-elements are needed (which allows a good accuracy of the plastic behaviour during the crack growth with no remeshing and no interpolation field) [3]. Concerning the strategy of enrichment of the element containing the crack tip, an appropriate enrichment basis for fatigue crack growth simulation with confined plasticity is used [5]:

$$r^{1/(n+1)} \left\{ \sin\frac{\theta}{2}, \cos\frac{\theta}{2}, \sin\frac{\theta}{2}\sin\theta, \cos\frac{\theta}{2}\sin\theta, \sin\frac{\theta}{2}\sin 3\theta, \cos\frac{\theta}{2}\sin 3\theta \right\}, \quad (5)$$

where n is linked to the hardening of the elastic-plastic model. Indeed, the plastic asymptotic HRR fields can be well represented by the following basis (non-linear elasticity). In this respect, such enrichment can be well suited for elasto-plasticity when the load is increasing. Furthermore, numerical simulations show that the following basis gives accurate results for specimens

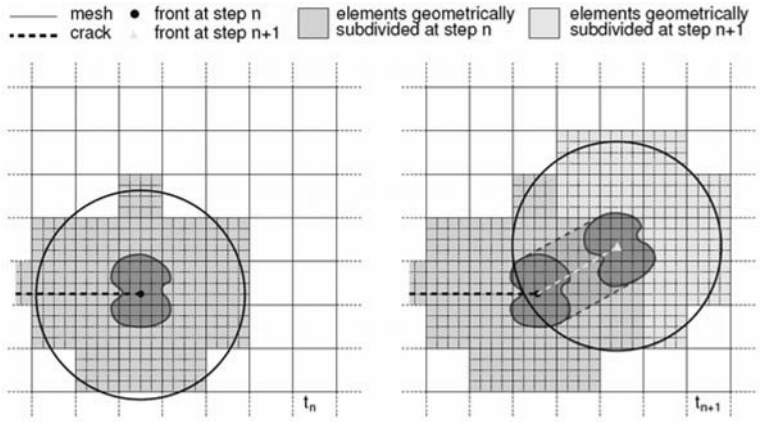


Fig. 12. Evolution of the sub-elements between the two configurations

submitted to loading and unloading even with coarse mesh for a wide range of elastic-plastic hardening [5].

In order to take into account contact and friction between the crack faces, interface elements are implemented along the crack with their own displacement and traction fields (t and w). In this respect, a three fields Augmented Lagrangian formulation is introduced which allows to couple the interface elements with the displacement and the Cauchy stress fields (u and σ) [3]:

$$\int_\Omega \sigma_n^{(i)} : \varepsilon(u^*) d\Omega = \int_{\partial_2 \Omega} F_d . u^* dS + \int_\Gamma \Lambda_n^{(i)} . u^*|_\Gamma dS + \int_\Gamma \Lambda^* . (u_n^{(i)}|_\Gamma - w_n^{(i)}) dS$$
$$+ \int_\Gamma (t_n^{(i-1)} + \alpha w_n^{(i-1)}) . w^* dS - \int_\Gamma (\Lambda_n^{(i)} + \alpha w_n^{(i)}) . w^* dS$$
$$\forall (u^*, w^*) \in U_0, \forall \Lambda^* \in L_0, \tag{6}$$

where σ and ε are respectively the Cauchy stress and the strain fields; t and w respectively the load and displacement fields along the crack faces; F_d the prescribed load; u^*, w^* and Λ^* respectively the virtual displacement field in the bulk, the virtual displacement field on the crack faces and the virtual Lagrange multiplier field; α the penalty term (for the normal and the tangential problem) and Λ the Lagrange multiplier of the Augmented Lagrangian formulation. Finally, concerning the update of the local enrichment during propagation the following strategy is retained: in the aim to preserve the history of the thermodynamics variables, all enrichments are retained during crack growth [15]. As a consequence, the change of discretization is performed by imposing the new crack segment to be closed (new enrichments are initialized to zero; the new t and w must be compatible with u and σ). Furthermore, the new stress and internal variables are initialized with elastic conditions thanks to the moving subdividing zone (see Figure 13).

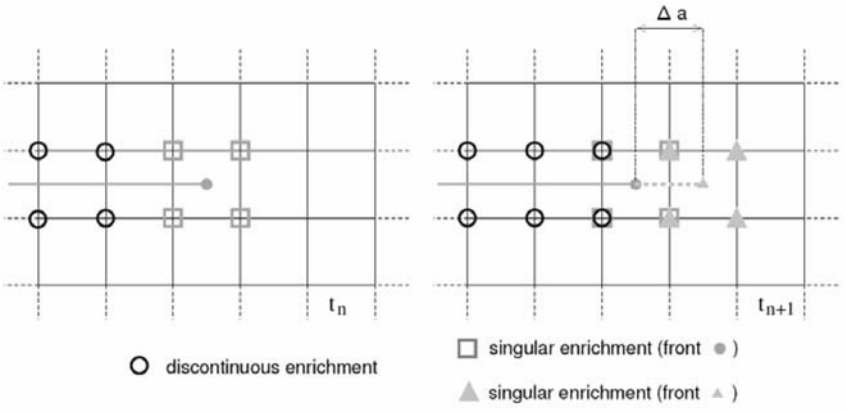


Fig. 13. Local enrichment and initialization strategy

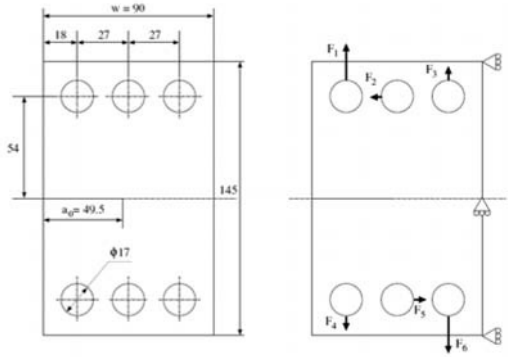


Fig. 14. Geometry and boundary conditions for the CTS specimen

4.2 Application to Fatigue Crack Growth with Confined Plasticity

One considers as a first example a Compact Tension Shear (CTS) specimen, submitted to mode I fatigue crack growth with interspersed overload. A numerical and experimental investigation of that specimen is fully presented in [9]. In this example, one wants to compare the results obtained by a classical FE calculation presented in [9] with the proposed method. In Figure 14, the geometry and boundary conditions are presented, and one also illustrates the mesh of the specimen with a zoom around the crack tip for the FE simulations and the X-FEM simulations.

The material is chosen similar to the one in [9]: 2.1011 Pa for the Young's modulus, 0.3 for the Poisson's ratio, 200 MPa for the yield strength, 534 MPa for the hardening coefficient, 0.27 for the hardening exponent n, and 90 MPa.m$^{1/2}$ for the critical mode I SIF.

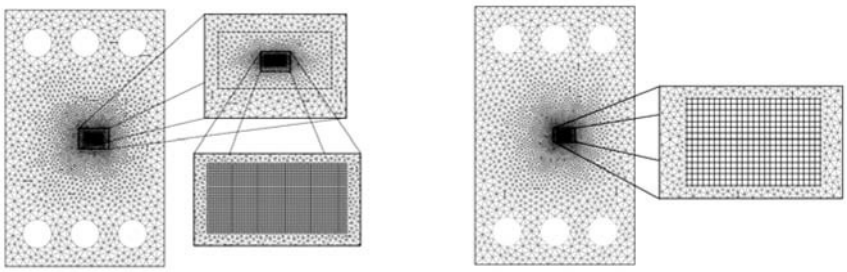


Fig. 15. Mesh of the CTS specimen with zoom around the crack tip

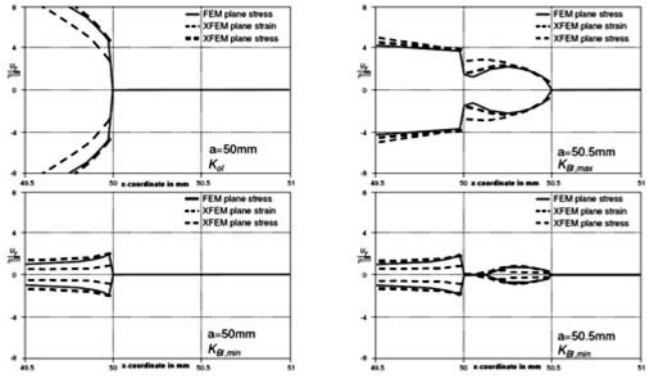


Fig. 16. Comparison of the vertical displacement on the crack faces between FEM and X-FEM before and after the propagation

In this comparison, the X-FEM mesh as been chosen 10 times coarser than the FEM mesh. Furthermore, a standard von Mises plasticity with isotropic hardening (Newton+radial return scheme) has been considered here [3]. The specimen is submitted to a total crack growth of $\Delta a = 0.5$ mm (20 steps in FEM, 2 steps in X-FEM), then submitted to an overload with a ratio of 2.5 and then a total growth of delta = 0.5 mm. One can conclude on this example that, even with a 10 time coarser element length and a 10 time coarser time discretization, very good agreements can be obtained between FEM and X-FEM on the vertical displacement close to the crack tip before and after the propagation of the crack. One can also notice the closure of the crack due to the overload and the confined plasticity (after the propagation).

As a second example, one considers a mode I Compact Tension (CT) specimen with a loading ratio of 0.1. The geometry and dimensions of the specimen are given in Figure 17. The material is chosen to be similar to the previous example. The specimen is submitted to a tension cyclic loading in order to have a stabilized stress state around the tip, then a growth of $\Delta a = 0.05$ mm is imposed at maximum load, and so on.

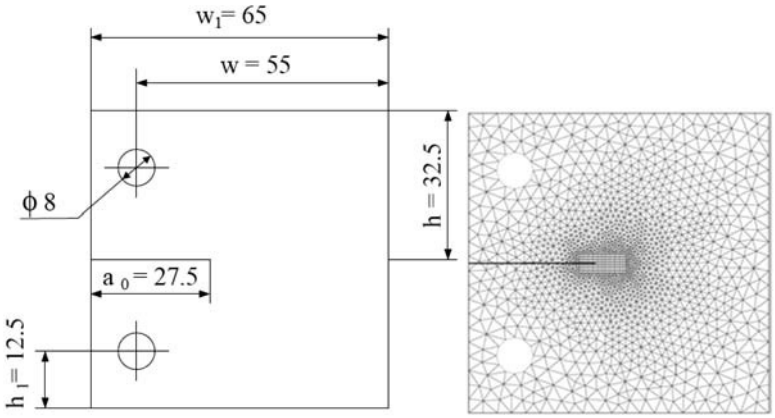


Fig. 17. Geometry and X-FEM mesh for the mode I CT specimen

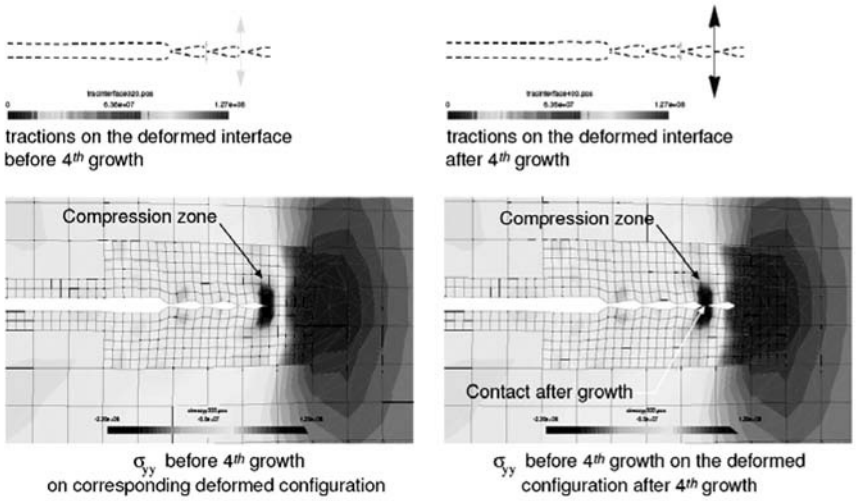


Fig. 18. Stress and interface results at minimum loads

In Figure 18, one illustrates the ability of the method to simulate closure effect close to the crack tip. In Figure 19, one illustrates stress-strain curves obtained on a Gauss point close to the crack tip during the propagation and the cycles of stabilization. One can notice that, due to the fact that no remeshing is done during the propagation of the crack, it is very easy to follow the stress-strain evolution of the material in a fixed point, which can not be possible with the same accuracy if you proceed to remeshing and interpolations.

Concerning the time dependence of the problem, this case can be considered as a two scale approach: a fixed "coarse" scale linked to the geometry (structure and crack) and a "fine" scale linked to the non-linear properties

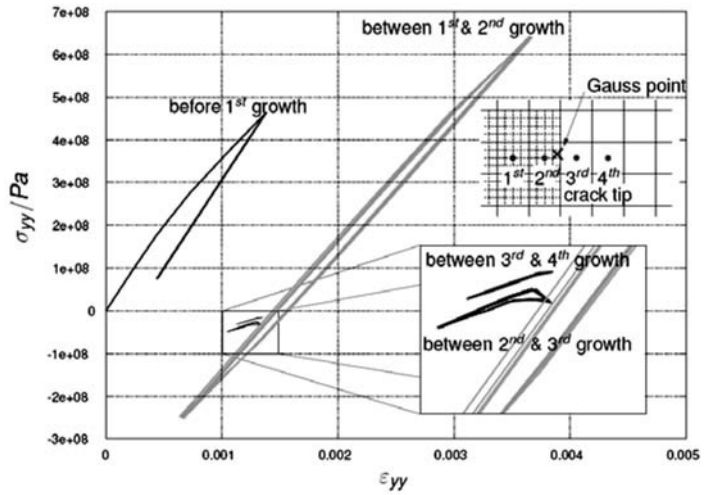


Fig. 19. Stress-strain evolution during crack growth

of the material (confined plasticity, contact and friction between the crack faces) which follows the crack tip during the propagation. In this respect, no remeshing is needed; however, new integration points are needed (it consists to initialize the new internal variables to zero according to the linear elastic behaviour outside the crack tip plastic zone), and new interface elements are needed.

Numerical simulations show that the proposed basis of enrichment can give accurate results for specimens submitted to loading and unloading even with coarse meshes for a wide range of hardening. However, because of conditioning issues, the crack extension can not be small compared to the element size (discrete successive plastic zone). For cyclic loadings with elastic-plastic crack growth, the question of the optimal enrichment basis is still open. One can also notice that the effects of non-singular terms of the asymptotic fields like T-stress are not taken into account. Furthermore, numerical issues suck as locking, plastic incompressibility and convergence rate need specific studies.

5 Conclusions

The propagation of a semi-elliptical crack in the bulk of an ultrafine-grained Al–Li alloy has been investigated using synchrotron radiation X-ray micro-tomography. In this material, the studied crack, despite its small dimension, can be considered as a "microstructurally long" and described in the frame of the linear elastic fracture mechanics. The extended finite element method was used to calculate the stress intensity factors along the crack front taking into account the three-dimensional geometry extracted from the tomographic

images. For the same nominal value of the stress intensity factor range, crack propagation was faster in the bulk than at the surface. The observed anisotropy is attributed to the variation of the closure stress along the crack front between surface and bulk. The experimentally observed fatigue crack propagation is compared to numerical simulations. Good agreement is found when a linear variation of closure stress along the crack front is taken into account in the "3D crack propagation law" used for the simulation. In a second example, it is shown that even with elastic-plastic behaviour coupled with contact and friction, non-remeshing property can be preserved for instance for 2D mixed-mode plastic fatigue crack growth. The main idea consists in using a specific nonlinear enrichment basis which allows to take into account the asymptotic behaviour around the crack tip. Because of the treatment of multiple nonlinearities (plasticity, contact), the numerical integration is adapted in order to properly integrate the enrichment basis and to have a fine knowledge of the stress state around the crack tip. Second, the propagation strategy consists in keeping the history of the variables such as stresses and plastic strains. Finally, an original strategy was designed in order to avoid projection of stresses and internal variables as the crack evolves to ensure the reliability of the method. In the presented examples, the method shows its ability to model the physical phenomena that are present in fatigue crack growth.

References

1. Moës N, Gravouil A, Belytschko T (2002) Non-planar 3D crack growth by the extended finite element and level sets, Part I : Mechanical model, International Journal for Numerical Methods in Engineering 53:2549–2568.
2. Gravouil A, Moës N, Belytschko T (2002) Non-planar 3D crack growth by the extended finite element and level sets, Part II : Level set update, International Journal for Numerical Methods in Engineering 53:2569–2586.
3. Elguedj T, Gravouil A, Combescure A (2006) A mixed Augmented Lagrangian eXtended Finite Element Method for modelling elastic-plastic fatigue crack growth with frictional contact. International Journal for Numerical Methods in Engineering, in press.
4. Ferri E, Buffière J-Y, Ludwig W, Gravouil A, Edwards L (2006) Three-dimensional study of fatigue crack propagation of a semi-elliptical crack: 3D crack visualization using X-ray microtomography and numerical simulation of crack propagation using the Extended Finite Element method, Acta Materialia 54:1111–1122.
5. Elguedj T, Gravouil A, Combescure A (2005) Appropriate extended functions for X-FEM simulation of plastic fracture mechanics, Computers Methods in Applied Mechanics and Engineering 195:501–515.
6. Elber W, Engineering Fracture Mechanics (1970) 2:37–45.
7. Venkateswara R, Ritchie R (1991) Metall. Trans. A 22:191–202.
8. Pitcher P, Bushby R, Vine W, Smith A, Tarrant A (2001) Mater. Sci. Technol. 17:807–814.

9. Sander M, Richard H (2005) Finite element analysis of fatigue crack growth with interspersed mode I and mixed mode overloads, International Journal of Fatigue 27:905–913.
10. Miller K (1982) Fatigue Eng. Mater. Struct. 5:223–232.
11. Sethian J-A (1996) Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Material Science. Cambridge University Press, Cambridge, UK.
12. Sukumar N, Moës N, Moran B, Belytschko T (2000) Extended Finite Element Method for Three-Dimensional Crack Modeling, International Journal for Numerical Methods in Engineering 48:1549–1570.
13. Wyart E, Duflot M, Coulon D, Martiny P, Pardoen T, Remacle J.-F, Lani F (2007) Substructuring FE-XFE approaches applied to three dimensional crack propagation, Journal of Computational and Applied Mathematics, in press.
14. Moës N, Dolbow J, Belytschko T (1999) A finite element method for crack growth without remeshing, International Journal for Numerical Methods in Engineering 46:131–150.
15. Réthoré J, Gravouil A, Combescure A (2005) An energy-conserving scheme for dynamic crack growth using the extended finite element method, International Journal for Numerical Methods in Engineering 63:631-659.

Accurate Simulation of Frictionless and Frictional Cohesive Crack Growth in Quasi-Brittle Materials Using XFEM

B.L. Karihaloo and Q.Z. Xiao

School of Engineering, Cardiff University, Queen's Buildings, The Parade, Newport Road, Cardiff CF24 3AA, UK; karihaloob@cardiff.ac.uk, qizhi.xiao@lusas.com

Summary. This paper discusses the crack tip asymptotic fields of frictionless and frictional cohesive cracks in quasi-brittle materials. This has been made possible after reformatting the cohesive-law into a special but universal polynomial. For accurate simulation of crack growth in quasi-brittle materials using the extended/generalized finite element method (XFEM), the leading term of the true displacement asymptotic field is used as the enrichment function at the tip of a cohesive crack. The opening component of the same field is also used as the initial guess opening profile of a newly extended cohesive segment. A statically admissible stress recovery (SAR) technique is used to recover the stresses at the crack tip. Finally, a pure mode I cohesive crack problem is analysed to demonstrate the characteristics of global responses and local fields obtained numerically by the XFEM.

Key words: Asymptotic expansion, cohesive crack, extended/generalized finite element method (XFEM), stress recovery.

1 Introduction

The cohesive zone (or crack) model of Hillerborg et al. [1] has been extensively used in the study of localisation and failure in quasi-brittle materials and structures. Elices et al. [2] have discussed its advantages and limitations. In the most widely used standard formulation of the model (see Figure 1), it is assumed that the stress–strain behaviour is isotropic linear elastic, and that the crack is initiated at a point where the maximum tensile principal stress σ_1 reaches the tensile strength f_t of the uncracked material, and that the crack is oriented normal to the direction of σ_1. An evolution law is also postulated for the monotonic mode I loading so that the cohesive stress is a unique function of the crack opening which, for concrete, decreases monotonically along the cohesive zone. The cohesive crack propagates when σ_1 at its tip reaches f_t. Although this standard formulation of the cohesive crack model is highly simplified, it is able to capture the essence of the fracture process in concrete specimens and structures [3].

Alain Combescure et al. (eds.), IUTAM Symposium on Discretization Methods for Evolving Discontinuities, 233–254.

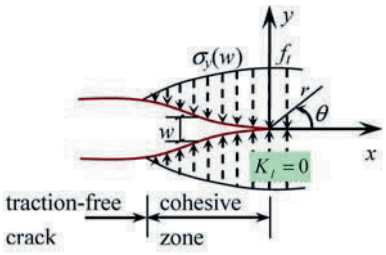


Fig. 1. A real traction-free crack terminating in a fracture process (cohesive) zone (FPZ) with residual stress transfer capacity $\sigma_y(w)$ whose faces close smoothly near its tip ($K_I = 0$). The material outside the FPZ is linear elastic, but within the FPZ is softening.

De Borst et al. [4] have given a concise overview of the various ways to numerically implement the cohesive zone methodology. They concluded that the extended/generalized finite element method (XFEM) [5–8] provides a proper representation of the discrete character of cohesive zone formulations avoiding any mesh bias. The XFEM enriches the standard local FE approximations with known information about the problem, such as a displacement discontinuity across a crack, the asymptotic solution at a crack tip, or a strain discontinuity across an interface, with the use of the partition of unity (PU). In contrast with the FEM, it avoids the use of meshes conforming with the discontinuity and adaptive remeshing as the discontinuity grows.

In the XFEM formulation of the cohesive zone model, Wells and Sluys [9] used the jump function as an enrichment function for the whole cohesive crack, hence the cohesive crack tip touches the element boundary. Moës and Belytschko [10] used the jump function for the part of the cohesive crack not adjacent to its tip, and a branch function adjacent to the tip. This approach can handle cases in which the cohesive crack tip does not touch the element boundary. However, the adopted branch function does not represent the true asymptotic nature of the displacement/stress field adjacent to the cohesive crack tip. Hansbo and Hansbo [11] modelled strong and weak discontinuities in solid mechanics by considering an element traversed by a discontinuity as a double element each half of which is used for the interpolation of one side of the discontinuity. This approach is difficult to use when the discontinuity ends within an element. Zi and Belytschko [12] enriched all cracked linear 3-node or quadratic 6-node triangular elements including the elements containing the crack tip by the sign function. Alfaiate et al. [13] embedded displacement jumps which do not need to be homogeneous within each FE. Mariani and Perego [14] introduced in a standard FE model a cubic displacement discontinuity, in order to reproduce the typical cusp-like shape of the process zone at the tip of a cohesive crack. However, this cubic function does not represent the true angular distribution of the displacement adjacent to the tip.

The lack of any work on the asymptotic fields at the tip of a cohesive crack is surprising considering the widespread use of cohesive crack models. This blank has been recently filled by Xiao and Karihaloo [15] who obtained universal asymptotic expansions at a cohesive crack tip, analogous to Williams expansions at a traction-free crack tip. Coulomb friction on the cohesive crack faces is also considered. The coefficients of the expansions of course depend nonlinearly on the softening-law and the boundary conditions. These universal expansions are valid for any normal cohesion-separation law (i.e. softening law) that can be expressed in a special polynomial form. They demonstrated that many commonly-used cohesion-separation laws, e.g. rectangular, linear, bilinear and exponentional, can be expressed in this special form.

In this paper, we will summarize the reformulated polynomial softening-law and the crack tip asymptotic fields of frictionless and frictional cohesive cracks in quasi-brittle materials given in [15]. We will use the leading term of the true displacement asymptotic field at the tip of a cohesive crack as the enrichment function in the XFEM. The opening component of the same field is also used as the initial guess opening profile of a newly extended cohesive segment. The tip of the growing cohesive crack does not have to terminate on an element boundary.

Rubinstein [16] has shown that relatively small errors in the determination of the crack path deflection angle can lead to a significant cumulative deviation of the crack path over a finite crack length. Therefore a reliable analysis of cohesive crack propagation requires an accurate knowledge of the crack tip field. In order to obtain accurate angular distribution of the stress adjacent to the crack tip, a statically admissible stress recovery (SAR) scheme [17–19] will be adopted. SAR uses basis functions, which meet the equilibrium equations within the domain and the local traction conditions on the boundary, and moving least squares (MLS) to fit the stresses at sampling (e.g., quadrature) points obtained by the XFEM. The most widely used stress recovery scheme (denoted as AVG) will also be used for comparison. AVG simply averages the stress values at each node evaluated from adjacent elements by bi-linear extrapolation from the Gauss points, and interpolates the averaged stresses using shape functions.

We will analyse a notched flexural specimen made of quasi-brittle materials with a bilinear law and show that the growth of a cohesive crack in a quasi-brittle material can be accurately predicted with a very coarse mesh which is not possible with the FEM. We will also demonstrate the characteristics of global responses and local fields obtained numerically by the XFEM.

2 Polynomial cohesive law for quasi-brittle materials

In order to obtain the separable asymptotic field at a cohesive crack tip (in terms of r and θ functions, r and θ are local polar coordinates centered at the tip, see Figure 1) in quasi-brittle materials like concrete, Xiao and Karihaloo

[15] reformulated the softening law into the following polynomial

$$\frac{\sigma}{f_t} = 1 + \sum_{i=1}^{5} \alpha_i \left(\frac{w}{w_c}\right)^{(2/3)i} - \left(1 + \sum_{i=1}^{5} \alpha_i\right) \left(\frac{w}{w_c}\right)^4 \tag{1}$$

or in normalized form

$$\hat{\sigma}_y = 1 + \sum_{i=1}^{5} \alpha_i \hat{w}^{(2/3)i} - \left(1 + \sum_{i=1}^{5} \alpha_i\right) \hat{w}^4, \tag{2}$$

where σ and f_t are the stress normal to the cohesive crack face and the uniaxial tensile strength, respectively; w and w_c are the opening displacement of the cohesive crack faces, and the critical opening displacement of the pre-existing macrocrack tip when it begins to grow; α_i, $i = 1 \ldots 5$, are fitting parameters. Relation (1) or (2) can represent a wide variety of softening laws.

For example, the widely used linear softening law

$$\frac{\sigma}{f_t} = 1 - \frac{w}{w_c} \tag{3}$$

can be represented by (1) with the only non-vanishing coefficients $\alpha_1 = -0.2612$, $\alpha_2 = -1.0215$, i.e.

$$\hat{\sigma}_y = 1 + \alpha_1 \hat{w}^{(2/3)i} + \alpha_2 \hat{w}^{4/3} - (1 - \alpha_1 + \alpha_2)\hat{w}^2. \tag{4}$$

The correlation coefficient is 1. In this study, the correlation coefficient is the Pearson's coefficient adopted by Mathcad. The linear law (3) is compared with (1) or (4) in Figure 2; they cannot be distinguished on the scale of the figure.

Similarly, for the widely used bilinear tension-softening law (Figure 3)

$$\hat{\sigma} = \begin{cases} 1 - (1 - \hat{f}_1)\dfrac{\hat{w}}{\hat{w}_1}, & 0 \le \hat{\sigma} \le \hat{f}_1, \\ \dfrac{\hat{f}_1}{1 - \hat{w}_1}(1 - \hat{w}), & \hat{f}_1 < \hat{\sigma} \le 1. \end{cases} \tag{5}$$

($\hat{f}_1 = f_1/f_t, \hat{w}_1 = w_1/w_c$), its two linear parts can be rewritten into two linear laws as shown in Figure 3. The first part can be written into (4) using a new definition of w_c as

$$w_c = \frac{w_1}{1 - \hat{f}_1}. \tag{6}$$

The second part can be written into (4) using a new definition of f_t as

$$f_t = \frac{f_1}{1 - \hat{w}_1}. \tag{7}$$

Cornelissen et al. [20] introduced the following exponential relation

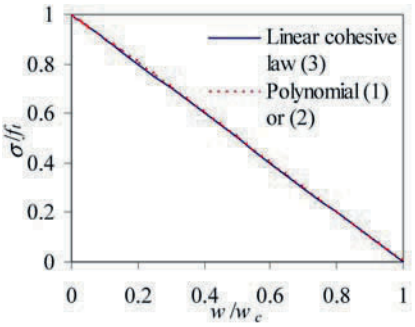


Fig. 2. Linear tension-softening law.

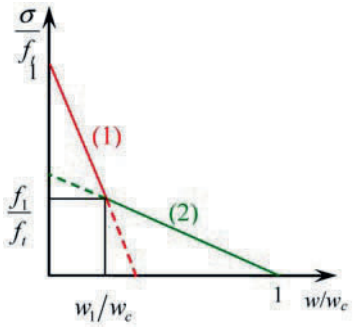


Fig. 3. Bilinear tension-softening law.

$$\frac{\sigma}{f_t} = f\left(\frac{w}{w_c}\right) - \frac{w}{w_c} f(1), \quad f\left(\frac{w}{w_c}\right) = \left[1 + \left(C_1 \frac{w}{w_c}\right)^3\right] e^{-C_2 (w/w_c)}. \tag{8}$$

C_1 and C_2 are fitting parameters. For a normal concrete (NC) with density 2370 kg/m^3, compressive strength $f_c = 47$ MPa, Young modulus $E = 39$ GPa, $f_t = 3.2$ MPa, $w_c = 160$ μm, specific fracture energy $G_F = 100$ J/m^2 (area under the tension-softening curve), $C_1 = 3$ and $C_2 = 6.93$. This diagram can also be fitted by (1) with $\alpha_1 = -0.872$, $\alpha_2 = -16.729$, $\alpha_3 = 67.818$, $\alpha_4 = -110.462$, and $\alpha_5 = 83.158$. The correlation coefficient is 1. They are compared in Figure 4, and cannot be distinguished on the scale of the figure.

Wecharatana [21] introduced the following softening relationship

$$\hat{\sigma}^m + \hat{w}^{2m} = 1, \tag{9}$$

$m = 0.27$ for concrete with compressive strength $f_c = 24$ MPa. In the range of $0 \le \hat{w} \le 0.6$, we can fit (9) using (1) with $\alpha_1 = -6.9495$, $\alpha_2 = 29.9794$, $\alpha_3 = -87.2663$, $\alpha_4 = 148.3647$, and $\alpha_5 = -128.84$. The correlation coefficient is 1. When $\hat{w} = 0.6$, $\hat{\sigma} = 0.005148$; when $\hat{w} > 0.6$, $\hat{\sigma}$ is negligibly small. As compared in Figure 5, they cannot be distinguished on the scale of the figure.

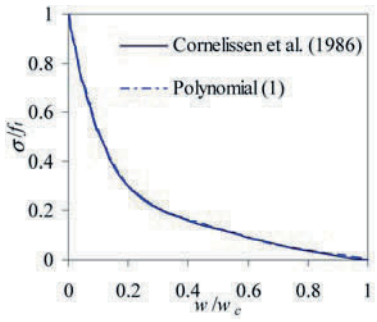


Fig. 4. Cohesive law (8).

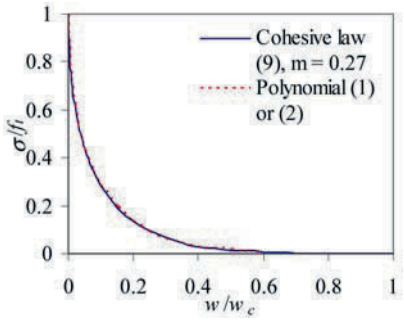


Fig. 5. Cohesive law (9).

3 Crack tip asymptotic fields of frictionless and frictional cohesive cracks in quasi-brittle materials

Muskhelishvili [22] showed that, for plane problems, the stresses and displacements in the Cartesian coordinate system (see Figure 1) can be expressed in terms of two analytic functions $\phi(z)$ and $\chi(z)$ of the complex variable $z = re^{i\theta}$

$$\begin{aligned} &\sigma_+ + \sigma_y = 2[\phi'(z) + \overline{\phi'(z)}], \\ &\sigma_y - \sigma_+ + 2i\tau_{xy} = 2[\bar{z}\phi''(z) + \chi''(z)], \\ &2\mu(u + i\nu) = \kappa\phi(z) - \overline{z\phi'(z)} - \overline{\chi'(z)}, \end{aligned} \tag{10}$$

where a prime denotes differentiation with respect to z and an overbar complex conjugate. In (10), $\mu = E/[2(1+\nu)]$ is the shear modulus; the Kolosov constant is $\kappa = 3 - 4\nu$ for plane strain or $\kappa = (3 - \nu)/(1 + \nu)$ for plane stress.

For a general plane mixed mode I + II problem, the complex functions $\phi(z)$ and $\chi(z)$ can be chosen as series of complex eigenvalue Goursat functions [23, 24]

$$\phi(z) = \sum_{n=0} A_n z^{\lambda_n} = \sum_{n=0} A_n r^{\lambda_n} e^{i\lambda_n \theta},$$

$$\chi(z) = \sum_{n=0} B_n z^{\lambda_n+1} = \sum_{n=0} B_n r^{\lambda_n+1} e^{i(\lambda_n+1)\theta}, \tag{11}$$

where the complex coefficients are $A_n = a_{1n} + ia_{2n}$ and $B_n = b_{1n} + ib_{2n}$. The eigenvalues λ_n and coefficients a_{1n}, a_{2n}, b_{1n} and b_{2n} are real.

The asymptotic analysis of the crack tip fields needs to satisfy the proper symmetry conditions along the line of extension of the cohesive crack, and boundary conditions on the cohesive crack faces.

If normal cohesive separation applies to the crack faces, relationship (1) needs to be satisfied over the cohesive zone. The stresses at the cohesive crack tip are non-singular (because the stress intensity factor $K_I = 0$). Moreover, the following conditions need to be satisfied:

(a) if the Coulomb friction on the crack face is considered

$$\sigma_y|_{\theta=\pi} = \sigma_y|_{\theta=-\pi} \neq 0, \quad \tau_{xy}|_{\theta=\pi} = \tau_{xy}|_{\theta=-\pi} = -\mu_f \sigma_y|_{\theta=\pm\pi} \neq 0, \tag{12}$$

where μ_f equals the positive or negative value of the coefficient of kinetic friction, which is assumed to be constant, depending on the relative sliding direction of the two crack faces: $\mu_f > 0$ when $\delta > 0$ and $\mu_f < 0$ when $\delta < 0$. When $\mu_f = 0$, the cohesive crack faces are frictionless.

(b) if the cohesive crack faces are in pure mode I condition

$$\sigma_y|_{\theta=\pi} = \sigma_y|_{\theta=-\pi} \neq 0, \quad \tau_{xy}|_{\theta=\pi} = \tau_{xy}|_{\theta=-\pi} = 0,$$
$$\tau_{xy}|_{\theta=0} = 0, \quad \text{and} \quad \nu|_{\theta=0} = 0. \tag{13}$$

The length of the process (cohesive) zone is determined by the condition $w \leq w_c$ in the normally opened part of the crack.

For a Coulomb frictional cohesive crack with normal cohesive separation (1), the complete asymptotic solutions are composed of two parts. The first part corresponds to integer eigenvalues

$$\lambda_n = n + 1, \quad b_{2n} = -\frac{n}{n+2} a_{2n} - \mu_f (a_{1n} + b_{1n}), \quad n = 0, 1, 2, \ldots, \tag{14}$$

or

$$\sigma_y|_{\theta=\pm\pi} = \sum_{n=0} (n+2)(n+1) r^n (a_{1n} + b_{1n}) \cos n\pi \quad \text{or}$$

$$\hat{\sigma}_y = \frac{\sigma_y|_{\theta=\pm\pi}}{f_t} = \sum_{n=0} c_n r^n = 1 + \sum_{n=1} c_n r^n, \tag{15}$$

where

$$c_n = \frac{(n+2)(n+1)(a_{1n} + b_{1n}) \cos n\pi}{f_t}, \quad c_0 = \frac{2(a_{10} + b_{10})}{f_t} = 1 \tag{16}$$

since $\sigma_y|_{\theta=\pm\pi} = f_t$ when $r \to 0$. These solutions have non-zero σ_y and τ_{xy} along the cohesive crack faces, but zero crack opening w and sliding δ. The second part of the asymptotic solutions corresponds to non-integer eigenvalues

$$\lambda_n = \frac{2n+3}{2}, \quad b_{1n} = -\frac{2n+1}{2n+5}a_{1n}, \quad b_{2n} = -a_{2n}, \quad n = 0, 1, 2, \ldots,$$

$$w = \sum_{n=0} \frac{r^{(2n+3)/2}}{\mu} \left[\left(\kappa + \frac{2n+3}{2} \right) a_{1n} + \frac{2n+5}{2} b_{1n} \right] \sin \frac{2n+3}{2}\pi \tag{17}$$

or

$$\hat{w} = \frac{w}{w_c} = \sum_{n=0} \bar{d}_n r^{(2n+3)/2}, \quad \bar{d}_n = \frac{\left[\left(\kappa + \frac{2n+3}{2} \right) a_{1n} + \frac{2n+5}{2} b_{1n} \right] \sin \frac{2n+3}{2}\pi}{\mu w_c}. \tag{18}$$

Consider the truncated $N+1$ terms of $\hat{w}$ (18), and denote $d_0 = \bar{d}_0$, $d_n = \bar{d}_n/d_0$ $(n > 1)$

$$\hat{w} = d_0 r^{3/2} \left(1 + \sum_{n=1}^{N} d_n r^n \right). \tag{19}$$

The expansion of $\hat{w}$ (19) raised to the power $2i/3$ is also truncated to $N+1$ terms, since these terms include only the truncated $N+1$ terms of $\hat{w}$. Hence

$$\hat{w}^{(2/3)i} = d_0^{(2/3)i} r^i \left(1 + \sum_{n=1}^{N} \beta_{in} r^n \right), \tag{20}$$

where

$$\beta_{in} = \frac{f_i^{(n)}(0)}{n!}, \quad f_i(r) = \left(1 + \sum_{n=1}^{N} d_n r^n \right)^{(2/3)i} \tag{21}$$

and $f_i^{(n)}(0)$ denotes the nth derivative at $r = 0$. The first five coefficients β_{in} are

$$\beta_{i1} = \frac{2}{3} i d_1,$$

$$\beta_{i2} = \frac{1}{3} i \left(\frac{2}{3} i - 1 \right) d_1^2 + \frac{2}{3} i d_2,$$

$$\beta_{i3} = \frac{1}{9} i \left(\frac{2}{3} i - 1 \right) \left(\frac{2}{3} i - 2 \right) d_1^3 + \frac{2}{3} i \left(\frac{2}{3} i - 1 \right) d_1 d_2 + \frac{2}{3} i d_3,$$

$$\beta_{i4} = \frac{1}{36} i \left(\frac{2}{3} i - 1 \right) \left(\frac{2}{3} i - 2 \right) \left(\frac{2}{3} i - 3 \right) d_1^4 + \frac{1}{3} i \left(\frac{2}{3} i - 1 \right) \left(\frac{2}{3} i - 2 \right) d_1^2 d_2$$
$$+ \frac{1}{3} i \left(\frac{2}{3} i - 1 \right) d_2^2 + \frac{2}{3} i \left(\frac{2}{3} i - 1 \right) d_1 d_3 + \frac{2}{3} i d_4.$$

$$\beta_{i5} = \frac{1}{180} i \left(\frac{2}{3} i - 1 \right) \left(\frac{2}{3} i - 2 \right) \left(\frac{2}{3} i - 3 \right) \left(\frac{2}{3} i - 4 \right) d_1^5$$
$$+ \frac{1}{9} i \left(\frac{2}{3} i - 1 \right) \left(\frac{2}{3} i - 2 \right) d_1^3 d_2 + \frac{1}{3} i \left(\frac{2}{3} i - 1 \right) \left(\frac{2}{3} i - 2 \right) d_1 d_2^2$$

$$+\frac{1}{3}i\left(\frac{2}{3}i-1\right)\left(\frac{2}{3}i-2\right)d_1^2d_3+\frac{2}{3}\left(\frac{2}{3}i-1\right)d_2d_3$$
$$+\frac{2}{3}i\left(\frac{2}{3}i-1\right)d_1d_4+\frac{2}{3}id_5. \tag{22}$$

If we choose $N=5$, then after satisfying the cohesive relationship (1) we have the following expressions for coefficients c_n in (16)

$$\begin{aligned}
c_1 &= \alpha_1 d_0^{2/3},\\
c_2 &= \alpha_2 d_0^{4/3}+\alpha_1 d_0^{2/3}\beta_{11},\\
c_3 &= \alpha_3 d_0^2+\alpha_1 d_0^{2/3}\beta_{12}+\alpha_2 d_0^{4/3}\beta_{21},\\
c_4 &= \alpha_4 d_0^{8/3}+\alpha_1 d_0^{2/3}\beta_{13}+\alpha_2 d_0^{4/3}\beta_{22}+\alpha_3 d_0^2\beta_{31},\\
c_5 &= \alpha_5 d_0^{10/3}+\alpha_1 d_0^{2/3}\beta_{14}+\alpha_2 d_0^{4/3}\beta_{23}+\alpha_3 d_0^2\beta_{32}+\alpha_4 d_0^{8/3}\beta_{41},\\
c_6 &= \alpha_1 d_0^{2/3}\beta_{15}+\alpha_2 d_0^{4/3}\beta_{24}+\alpha_3 d_0^2\beta_{33}+\alpha_4 d_0^{8/3}\beta_{42}\\
&\quad +\alpha_5 d_0^{10/3}\beta_{51}-\left(1+\sum_{i=1}^{5}\alpha_i\right)d_0^4,\\
c_7 &= \alpha_2 d_0^{4/3}\beta_{25}+\alpha_3 d_0^2\beta_{34}+\alpha_4 d_0^{8/3}\beta_{43}+\alpha_5 d_0^{10/3}\beta_{52}\\
&\quad -\left(1+\sum_{i=1}^{5}\alpha_i\right)d_0^4\beta_{61},\\
c_8 &= \alpha_3 d_0^2\beta_{35}+\alpha_4 d_0^{8/3}\beta_{44}+\alpha_5 d_0^{10/3}\beta_{53}-\left(1+\sum_{i=1}^{5}\alpha_i\right)d_0^4\beta_{62},\\
c_9 &= \alpha_4 d_0^{8/3}\beta_{45}+\alpha_5 d_0^{10/3}\beta_{54}-\left(1+\sum_{i=1}^{5}\alpha_i\right)d_0^4\beta_{63},\\
c_{10} &= \alpha_5 d_0^{10/3}\beta_{55}-\left(1+\sum_{i=1}^{5}\alpha_i\right)d_0^4\beta_{64},\\
c_{11} &= -\left(1+\sum_{i=1}^{5}\alpha_i\right)d_0^4\beta_{65}.
\end{aligned} \tag{23}$$

For non-integer eigenvalues (17), the coefficients a_{1n} and a_{2n} may be regarded as independent, so that coefficients b_{1n} are linearly dependent on a_{1n} and b_{2n} on a_{2n}. For integer eigenvalues (14), coefficients a_{1n} and a_{2n} may also be regarded as independent, so that coefficients b_{2n} now depend linearly on a_{1n}, a_{2n} and b_{1n}. However, the coefficients b_{1n} for integer eigenvalues will

depend both linearly on a_{1n} for integer eigenvalues and nonlinearly on a_{1n} for non-integer eigenvalues via (16), (18), (19), (21) and (23). The inherent nonlinear nature of the problem is reflected in these nonlinear relationships between the coefficients of the asymptotic fields.

Note that the above asymptotic solution cannot be reduced to a pure mode I cohesive crack (compare (12) and (13)), since along the line of extension of the crack, $\theta = 0$, the shear stress does not vanish ($\tau_{xy} \neq 0$). However, the procedures are similar; and the complete asymptotic fields have been given in [15]. The first term corresponding to a non-integer eigenvalue that gives a normal displacement dis- continuity over the cohesive crack faces is

$$u = \frac{r^{3/2}}{2\mu} a_1 \left[\left(\kappa + \frac{1}{2} \right) \cos \frac{3}{2}\theta - \frac{3}{2} \cos \frac{1}{2}\theta \right],$$
$$v = \frac{r^{3/2}}{2\mu} a_1 \left[\left(\kappa - \frac{1}{2} \right) \sin \frac{3}{2}\theta - \frac{3}{2} \sin \frac{1}{2}\theta \right]. \tag{24}$$

4 Simulation of cohesive crack growth in quasi-brittle materials using XFEM

To model the cohesive cracks in the XFEM, a standard local FE displacement approximation around the crack is enriched with discontinuous Heaviside functions along the crack faces behind the crack tip including the open traction-free part, and the crack tip asymptotic displacement fields at nodes surrounding the cohesive crack tip using the PU. The approximation of displacements for an element can be expressed in the following form

$$\begin{Bmatrix} u^h(x) \\ v^h(x) \end{Bmatrix} = \sum_{i \in I} \phi_i(x) \begin{Bmatrix} u_{0i} \\ v_{0i} \end{Bmatrix} + \sum_{j \in J \cap I} \phi_j(x) H(x) \begin{Bmatrix} b_{1j} \\ b_{2j} \end{Bmatrix}$$
$$+ \sum_{m \in M_k \cap I} \phi_m(x) \begin{Bmatrix} u_m^{\text{tip}\,k} \\ v_m^{\text{tip}\,k} \end{Bmatrix}, \tag{25}$$

where I is the set of all nodes in the element, (u_{0i}, v_{0i}) are the regular degrees of freedom at node i, ϕ_i is the FE shape function associated with node i, J is the subset of nodes whose support is intersected by the crack but do not cover any cohesive crack tips, the function $H(x)$ is the Heaviside function centred on the crack discontinuity, and (b_{1j}, b_{2j}) are the corresponding additional degrees of freedom. M_k is the subset of nodes that are enriched around the cohesive crack tip k with the asymptotic displacements $u^{\text{tip}\,k}$ and $v^{\text{tip}\,k}$. $u_m^{\text{tip}\,k}$ and $v_m^{\text{tip}\,k}$ are enrichment functions adopted at node m with corresponding nodal parameters independent of other nodes.

Consider a domain Ω containing a crack Γ_c, as shown in Figure 6. The boundary Γ is composed of the segments Γ_u, Γ_t, and Γ_c. The part of the crack

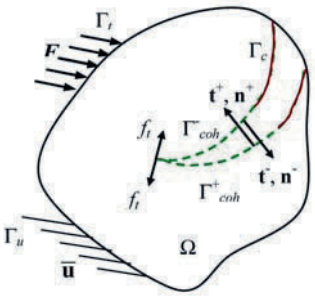


Fig. 6. Body with a crack Γ_c involving a cohesive zone Γ_{coh} subjected to prescribed boundary tractions $\mathbf{F}$ on Γ_t and displacements $\bar{\mathbf{u}}$ on Γ_u.

on which a softening law is active, i.e. the FPZ, is denoted by Γ_{coh}. Prescribed displacements $\bar{\mathbf{u}}$ are imposed on Γ_u, while tractions $\mathbf{F}$ are imposed on Γ_t. The crack surface Γ_c excluding Γ_{coh} is assumed to be traction-free. The cohesive tractions $\mathbf{t}^+$, $\mathbf{t}^-$ are imposed on the upper and lower surfaces of Γ_{coh}.

The virtual work equation without body forces is given by [9, 10]

$$\int_{\Omega} \sigma(\mathbf{u}) : \nabla^s \mathbf{v} d\Omega + \int_{\Gamma_{\text{coh}}} \mathbf{t} \cdot \mathbf{w}(\mathbf{v}) d\Gamma = \int_{\Gamma_t} \mathbf{F} \cdot \mathbf{v} d\Gamma, \tag{26}$$

where σ is the Cauchy stress, and ∇^s denotes the symmetric part of the gradient operator. The displacements $\mathbf{u}$ must belong to the space U of kinematically admissible displacement fields

$$\mathbf{u} \in U = \{\mathbf{u} : \mathbf{u} = \bar{\mathbf{u}} \text{ on } \Gamma_u, \mathbf{u} \text{ discontinuous on } \Gamma_c\}, \tag{27}$$

where the space U is related to the regularity of the solution. The test function $\mathbf{v}$ must belong to the space V defined by

$$\mathbf{v} \in V = \{\mathbf{v} : \mathbf{v} = \bar{\mathbf{v}} \text{ on } \Gamma_u, \mathbf{v} \text{ discontinuous on } \Gamma_c\}. \tag{28}$$

The cohesive tractions $\mathbf{t} = \mathbf{t}^+ = -\mathbf{t}^-$ and the separation $\mathbf{w}(\mathbf{v}) = \mathbf{v}^- - \mathbf{v}^+$ are related by a softening law on Γ_{coh}.

For a softening law composed of linear segments, equation (26) can be solved using a secant modulus iteration scheme when the length of the cohesive zone and external loads are given. The term secant modulus here refers to the fact that the stiffness matrix of XFEM obtained from (26) relates directly the total displacements to loads. Displacements (24) are used as the crack tip enrichment function in (25) for a mode I cohesive crack. The unknown coefficient a_1 depends on the softening law, and boundary and load conditions. It is considered as additional degrees of freedom at relevant enrichment nodes in XFEM. Details of the implementation of XFEM, SAR, and simulation of cohesive crack growth can be found in [19]. The flowchart is as follows:

Prescribe information on mesh and any traction-free crack, if it exists.
Carry out *initial linear analysis* to judge at which point σ_1 (recovered by, e.g. SAR) first reaches f_t. The corresponding load is used as the initial load for the next iteration.
Loop over *cohesive zone increments*
A cohesive segment of given length (user defined) is added at the location (i.e. at the front of a pre-existing traction-free or cohesive crack) where σ_1 (recovered by, e.g. SAR) reaches f_t. The initial guess opening for the first increment is set to be zero. Otherwise, the guess opening profile of the newly added cohesive segment is assumed to be of $cr^{3/2}$ type, where coefficient c is determined by the opening displacement at the closest integration point of the previous cohesive crack seg- ment.
The assumed load factor λ^1 for first load iteration is the converged value of the previous increment.
Iteration for *balance load*
Secant modulus **iteration** for *solving the displacement field using* (26); for a given external load and cohesive zone
If a linear softening law is used, the problem is linear and it is not necessary to check the convergence, **exit** *iteration*.
When a nonlinear softening law is used, if the ratio of the increment of the crack opening displacement in the current iteration to the total opening dis- placement (we just consider the part whose opening is larger than 0.1 of the maximum opening) $\leq 10^{-4}$, the solution is deemed to have converged and **exit** *secant modulus iteration*.
End *secant modulus* **iteration**
The stresses at the tip of the cohesive crack are calculated using AVG or SAR. If σ_1 at the cohesive crack tip satisfies $|f^{(k-1)}/f_t| = |(\sigma_1 - f_t)/f_t| < 10^{-3}$, **exit** *iteration for balance load*.
If $|f^{(k-1)}/f_t| \geq 10^{-3}$, compute the load factor of the next iteration step (k) using a *secant method*:
$k = 2$: $\lambda^{(2)} = \lambda^{(1)}(1.1 - 0.1\sigma_1/f_t), \quad f^{(1)} = \sigma_1 - f_t;$
$k > 2$: $f^{(k-1)} = \sigma_1 - f_t$
$\lambda^{(k)} = \lambda^{(k-1)} - (\lambda^{(k-1)} - \lambda^{(k-2)})f^{(k-1)}/(f^{(k-1)} - f^{(k-2)})$
If $\left|\frac{\lambda^{(k)} - \lambda^{(k-1)}}{\lambda^{(k-1)}}\right| > 0.1, \lambda^{(k)} = \lambda^{(k-1)}\left(1 + 0.05\frac{\lambda^{(k)} - \lambda^{(k-1)}}{|\lambda^{(k)} - \lambda^{(k-1)}|}\right)$
End iteration for *balance load*
If the last cohesive segment away from the tip has an opening at its central point $w > w_c$, it is removed from the cohesive zone and treated as a traction-free segment. The opening profile of the remaining part of the cohesive zone is used as the initial guess opening profile and the iterations for the new balance load are started for the modified cohesive zone.
If $w < w_c$, the balance load has been found for the *current cohesive increment*.
Output converged results.
If stop criterion has met, **exit** *cohesive zone increment*.
End loop over *cohesive zone increments*

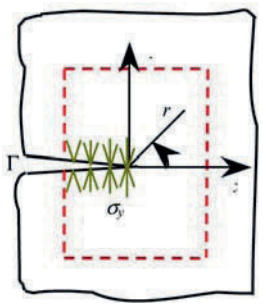


Fig. 7. Illustration of the cohesive crack problem.

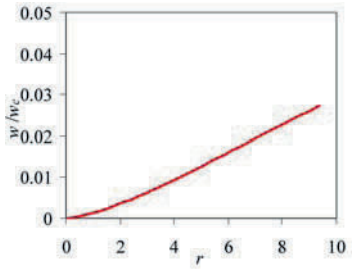


Fig. 8. Opening profile of the cohesive crack.

5 Illustration of the mode I cohesive crack tip fields

In order to visualise the cohesive crack tip fields derived in Section 3, it is necessary to determine the unknown coefficients by fitting the numerically computed crack tip fields with the theoretically obtained fields. This requires a sophisticated optimization scheme. For the present purpose of illustration, we will solve a mode I cohesive crack problem shown in Figure 7 using the obtained asymptotic fields and the numerically computed opening profile of the cohesive crack.

We consider the subdomain bounded by broken lines in Figure 7, and assume the length of the cohesive crack is 10.5 mm. The opening profile of the cohesive crack (Figure 8) adopts the computed results of a three point bend beam with a very small initial crack of length 0.1 mm at the bottom midpoint of the beam (Figure 9) at the loading stage with total cohesive crack = 31.5 mm and load/$(f_t bt)$ = 0.228 [15]. The geometrical parameters are b = 150 mm, $l = 4b$, $t = b$ (t is the specimen thickness in the out-of-plane direction). A linear softening law and a state of plane strain condition are considered. The material properties are E = 36.5 GPa, ν = 0.1, f_t = 3.19 MPa, G_F = 50 Nm/m^2. The dimensions for force and length are N and mm, respectively.

The displacements and stresses along the circle with radius r = 0.5 mm surrounding the crack tip and the line of extension of the cohesive crack are

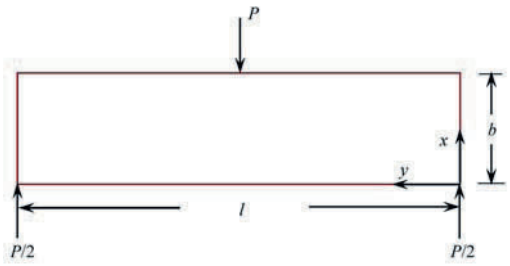


Fig. 9. An three-point bend beam (TPB).

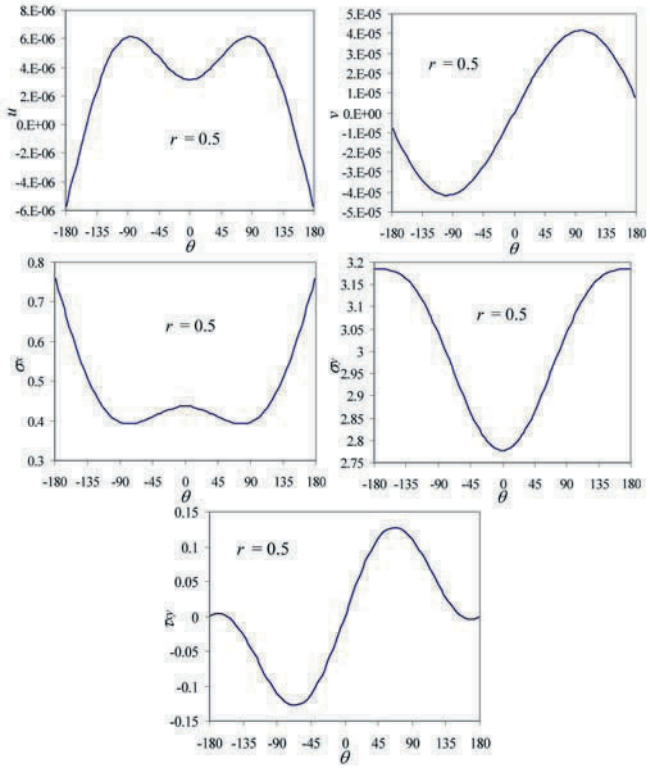


Fig. 10. Displacements and stresses along the circle $r = 0.5$.

plotted in Figures 10 and 11. As expected, the stress σ_y at the cohesive crack tip is equal to ft, and no stress at any other locations reaches f_t.

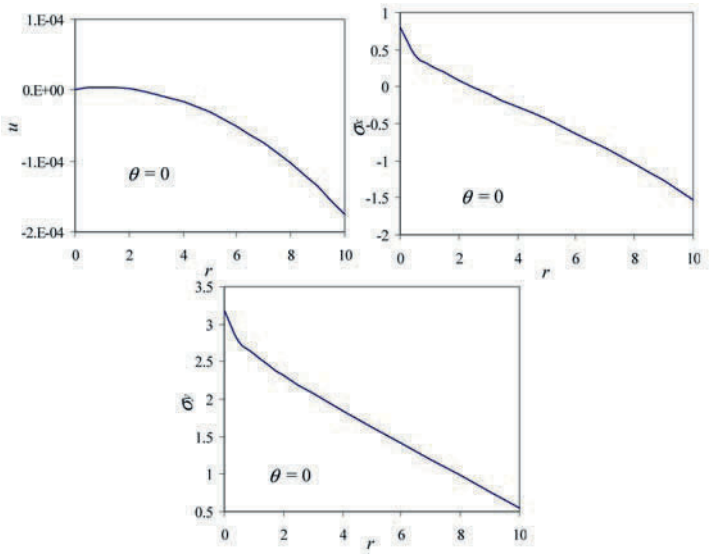


Fig. 11. Non-vanishing displacements and stresses along the line of extension of the cohesive crack ($\theta = 0$).

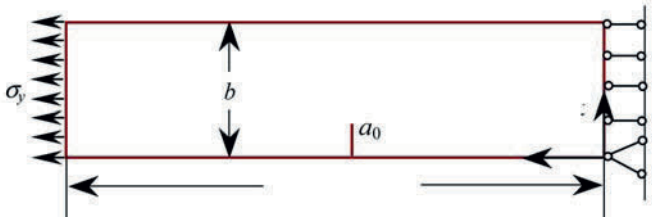


Fig. 12. An edge-cracked plate under uniaxial tension.

6 Numerical example

In this section, we will analyse an edge cracked plate under uniaxial tension (Figure 12). The geometrical parameters are $b = 150$ mm, $l = 4b$, $t = b$. A state of plane strain condition is considered. The initial edge crack has a length of $a_0 = 15.1$ mm. The properties $E = 36.9$ GPa, $\nu = 0.2$, and a bilinear softening law (5) with $f_t = 3.14$ MPa, $f_1 = 0.455$ MPa, $w_c = 0.279$ mm, $w_1 = 0.0373$ mm, $G_F = 122$ Nm/m^2 obtained by Abdalla and Karihaloo [25] for a real normal strength concrete are used.

Two meshes, as shown in Figure 13, are used in the analysis. The coarser mesh consists of $50 \times 100 = 5000$ rectangular elements, giving a total of 5151 nodes. The finer mesh consists of $150 \times 120 = 18000$ rectangular elements, giving a total of 18271 nodes. Both meshes are uniformly divided in x-direction. For the coarser mesh, the central 50 layers of elements have an identical height

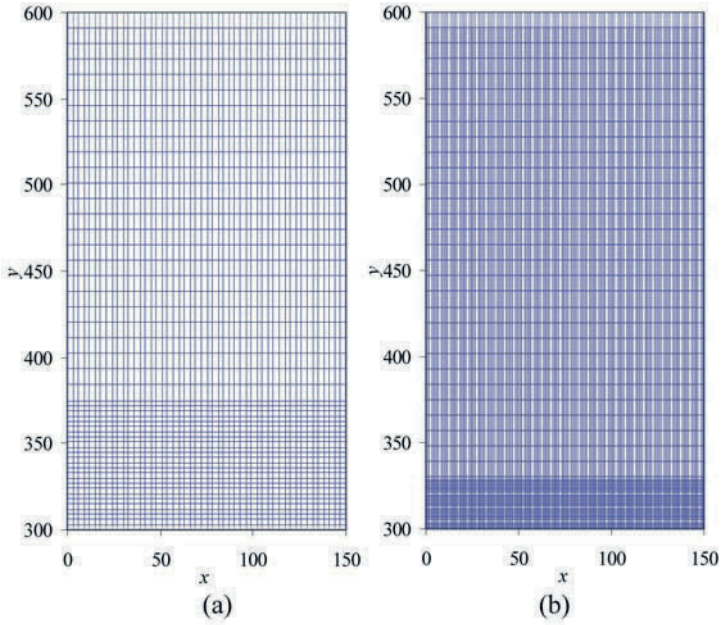


Fig. 13. Coarse (a) and fine (b) mesh for the left half of the specimen.

(y-direction) of 3mm; the remaining elements have an identical height of 9 mm. Therefore, elements in the central zone are 3×3 mm^2 squares. For the finer mesh, the central 60 layers of elements have an identical height of 1 mm; the remaining elements have an identical height of 9 mm. Therefore elements in the central zone are 1×1 mm^2 squares. We will study global responses including the load–deformation behaviour and the evolution of the cohesive zone, as well as local properties including the opening profile and distribution of the cohesive stress in the cohesive zone, and displacements and stresses along the circle $r = 3.5$ mm (the second layer of elements in the coarser mesh, or the fourth layer of elements in the finer mesh) around the crack tip and along the line of extension of the crack. The intention of using two meshes is to study the mesh size sensitivity of the global responses as well as of the crack tip fields. Furthermore, the results from the finer mesh will be used as a reference solution, since no analytical solutions are available. The conventional 4-node bilinear isoparametric Q4 elements are used as background elements. The potential fracture locus coincides with the specimen's axis of symmetry. The crack is modelled by enriching the nodes on the crack faces with jump without the double nodes that are used in the traditional FEM.

The first layer of nodes surrounding the cohesive crack tip are enriched with (24). The opening component of the same field is also used as the initial opening profile of a newly extended cohesive segment in the simulation of cohesive crack propagation.

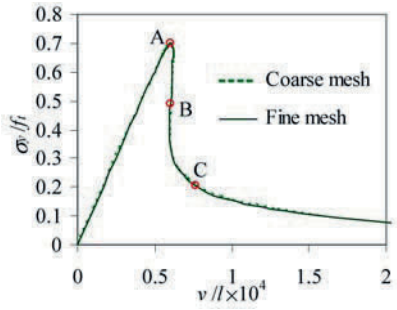


Fig. 14. The non-dimensional tensile stress-central extension curves of the uni-axial single-edge notched tension plate.

Similar to Moës and Belytschko [10], x-direction of nodes with coordinates $(0, 0)$ and $(0, 600$ mm) and y-direction of the node with coordinates (150 mm, 300 mm) are constrained; the load is distributed over a length of 6 mm for the coarse mesh and 2 mm for the fine mesh (two elements).

In the simulation, the first increment of the cohesive crack is 4.4 mm, then the cohesive crack propagates by a segment of length 3 mm after each step in the coarser mesh, and by three segments of length 1 mm each in the finer mesh.

The stresses at the tip of the cohesive crack recovered by SAR are used to judge whether or not the tip will propagate.

The stresses for the finer mesh recovered by AVG and SAR are identical along the circle $r = 3.5$ mm and the line of extension of the crack. Therefore, all stress results presented below are for the coarser mesh, unless otherwise mentioned.

The dimensions of the displacement and length parameters are in mm, and those of the stresses in MPa.

Our main concern is the accuracy of the angular distribution of crack tip displacements and stresses. However, it cannot be adequately measured by the widely used norms. It is also not suitable to use errors at particular locations. For instance, if the relative error of the maximum value of the displacement or stress is used, the error may be small but the overall angular distribution can still be very poor. On the other hand, if the maximum error in the angular direction for a given r is used, the relative error at locations where the exact values are very small can still be very large although the angular distribution is very accurate. Hence, we do not use any error measures. Instead, we compare the results directly on properly scaled plots.

The variation of the tensile stress σ_y and displacement v in y-direction at the central point of the loading edge is shown in Figure 14. Obviously, it shows no mesh dependency. The evolution of the size of the cohesive zone corresponding to the two branches of the bilinear softening law (5) is shown in Figure 15. It also reveals very weak mesh dependence. The opening profile

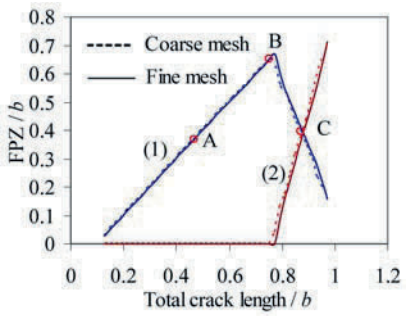


Fig. 15. Evolution of the cohesive zone size as the cohesive tip travels through the plate. (1) and (2) correspond to the first and second branches of the bilinear softening diagram (Figure 3).

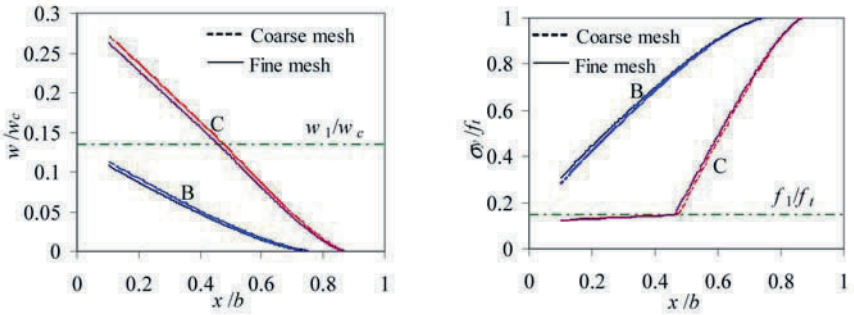


Fig. 16. Opening profile and distribution of cohesive stresses in the FPZ at typical loading stages (Figure 14).

and distribution of cohesive stresses in the cohesive zone are shown in Figure 16, and show some weak mesh dependence. Non-vanishing displacements and stresses along the circle $r = 3.5$ mm, and the line of extension of the crack at loading stages A are plotted in Figure 17. The displacements on $r = 3.5$ mm from the coarser mesh agree quite well with the finer mesh. The stresses for the coarser mesh obtained by SAR are more accurate than the direct differentiation and AVG, and agree quite well with the finer mesh. Along the line of extension of the crack, the displacement u in x-direction is very sensitive to the mesh size, however the remaining non-vanishing displacement and stress components are not. The agreement between AVG, SAR and the finer mesh is generally very good, however, SAR is better than AVG close to the cohesive crack tip.

Note that close to the cohesive crack tip, the maximum values of the main stresses σ_x and σ_y occur on the cohesive crack face. This feature of the stress distribution has also been noticed by Planas et al. [26].

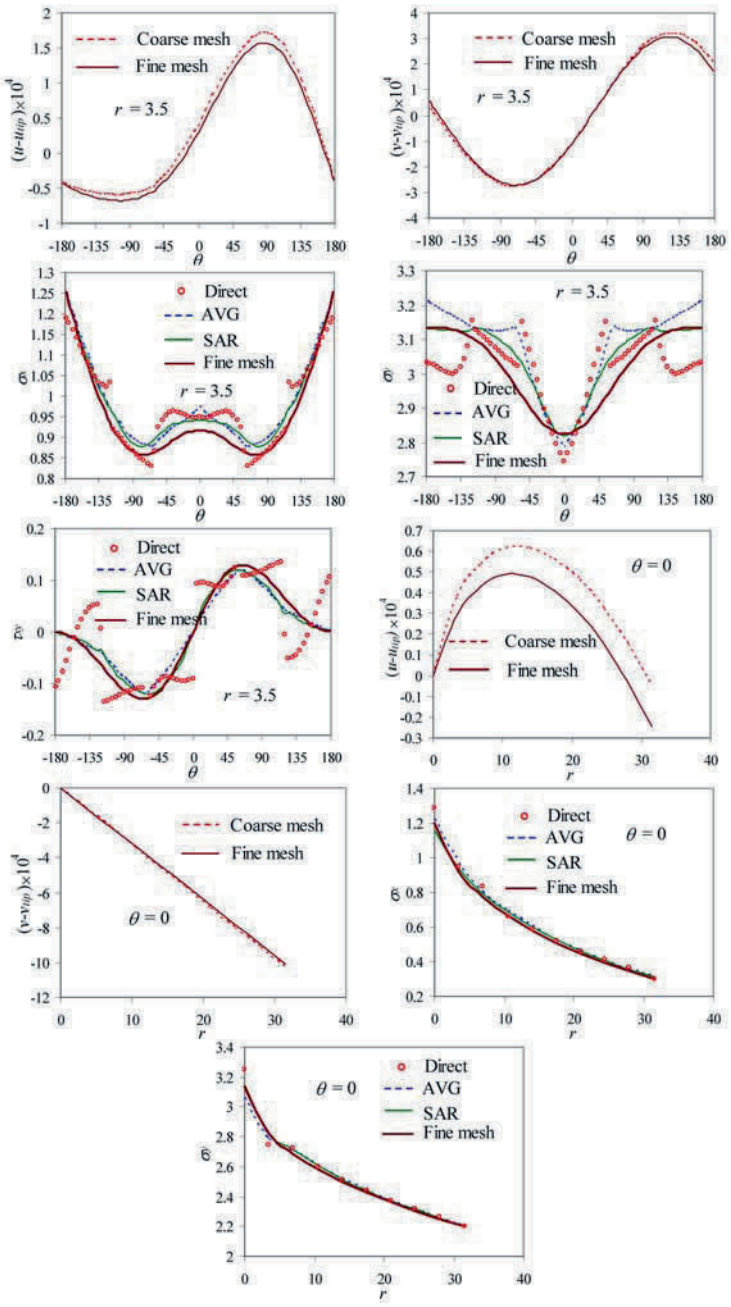


Fig. 17. Non-vanishing displacements and stresses along the circle $r = 3.5$ mm, and the line of extension of the crack at loading stage A (Figure 14).

The use of the enrichment function in the XFEM has been noticed to deteriorate the condition of the discrete system [6]. This issue can be alleviated by methods introduced in [6] and [27], or by using high performance solvers with partial pivoting such as routines MA27 and MA47 in Harwell Subroutine Library [28], and does not affect global responses or the stresses and the strain energy. However, it may affect significantly the displacements adjacent to the tip of the cohesive crack even alleviating methods are adopted, so that a direct comparison of the coarse and fine mesh patterns (Figure 13) adopted in the analysis is not possible. Hence, we compare instead the displacements relative to the tip of the cohesive crack.

7 Discussion and conclusions

The complete asymptotic expansions for frictionless and frictional cohesive cracks are analogous to the Williams expansions in brittle solids. They are valid for many commonly used separation laws, e.g. rectangular, linear, bi-linear, exponential, etc with or without Coulomb friction on crack faces. In numerical simulation of cohesive crack propagation, the asymptotic expansions can be used together with these separation laws. If relation (1) is used directly with the Newton–Raphson method, the tangential stiffness on the interface is infinite when $w = 0$. However, a large number may be used in stead as in the penalty function method.

Global responses like load–deformation curves, and evolution of the cohesive zone, are not sensitive to the size of the adopted mesh in the XFEM. Local properties like the opening profile and distribution of the cohesive stresses in the cohesive zone, especially crack tip displacements and stresses are more sensitive to the mesh size.

Although no singularity exists at the tip of a cohesive crack, the stresses obtained by direct differentiation of the displacements are not accurate, and cannot be used to predict accurately the growth of the tip. The SAR, which satisfies exactly the non-homogeneous cohesive tractions, gives more accurate results than the widely used AVG adjacent to the crack tip. If high accuracy of the local fields is required (e.g. these fields are use to judge the growth of the tip), a relatively fine mesh needs to be used together with accurate stress recovery methods, such as SAR.

For mode I cohesive cracks, the stresses recovered by SAR are highly accurate along the line of extension of the cohesive crack, and can be used directly to predict the growth of the crack. For general mixed-mode cracks, further studies are required on whether the SAR stresses at the tip, or on a small circle enclosing a few (e.g. one or two) layers elements around the tip should be used to predict the growth of the crack.

Acknowledgement

Financial support from the European Commission KMM-Network of Excellence is gratefully acknowledged.

References

1. Hillerborg A, Modeer E, Petersson PE (1976) Cement Concrete Res 6: 773–781
2. Elices M, Guinea GV, Gómez J, Planas J (2002) Engrg Fract Mech 69: 137–163
3. Karihaloo BL (1995) Fracture mechanics and structural concrete. Addison Wesley Longman, UK
4. de Borst R, Gutierrez MA, Wells GN, Remmers JJC, Askes H (2004) Int J Numer Meth Engrg 60: 289–315
5. Moës N, Dolbow J, Belytschko T (1999) Int J Numer Meth Engrg 46: 131–150
6. Strouboulis T, Copps K, Babuska I (2001) Comput Meth Appl Mech Engrg 190: 4081–4193
7. Karihaloo BL, Xiao QZ (2003) Comput Struct 81: 119–129
8. Xiao QZ, Karihaloo BL (2005) Recent developments of the extended/generalized FEM and a comparison with the FEM. In: Wu XP (ed) Developments and applications of solid mechanics (Proceedings of the Symposium on Prof M. G. Huang's 90th Birthday). Press of University of Science and Technology of China, Hefei, China, pp. 303–324
9. Wells GN, Sluys LJ (2001) Int J Numer Meth Engrg 50: 2667–2682
10. Moës N, Belytschko T (2002) Engrg Fract Mech 69: 813–833
11. Hansbo A, Hansbo P (2004) Comput Meth Appl Mech Engrg 193: 3523–3540
12. Zi G, Belytschko T (2003) Int J Numer Meth Engrg 57: 2221–2240
13. Alfaiate J, Simone A, Sluys LJ (2003) Int J Solids Struct 40: 5799–5817
14. Mariani S, Perego U (2003) Int J Numer Meth Engrg 58: 103–126
15. Xiao QZ, Karihaloo BL (2006) J Mech Mater Struct 1: 881–910
16. Rubinstein AA (2003) Int J Fract 119: L15–L20
17. Xiao QZ, Karihaloo BL (2004) Statically admissible stress recovery using the moving least squares technique. In: Topping BHV, Mota Soares CA (eds) Progress in computa- tional structures technology. Saxe-Coburg Publications, Stirling, Scotland, pp. 111–138
18. Xiao QZ, Karihaloo BL (2006) Int J Numer Meth Engrg 66: 1378–1410
19. Xiao QZ, Karihaloo BL, Liu XY (2007) Incremental-secant modulus iteration scheme and stress recovery for simulating cracking process in quasi-brittle materials using XFEM. Int J Numer Meth Engng 69: 2606–2635.
20. Cornelissen HAW, Hordijk DA, Reinhardt HW (1986) Heron 31: 45–56.
21. Wecharatana M (1990) Brittleness index of cementitious composite. In: Suprenant BA (ed) Serviceability and durability of construction materials, ASCE Publications, New York.
22. Muskhelishvili NI (1953) Some basic problems of mathematical theory of elasticity. Noordhoff, Holland.
23. Sih GC, Liebowitz H (1968) Mathematical theories of brittle fracture. In: Liebowitz H (ed) Fracture: An advanced treatise, Vol. II. Academic Press, New York, pp. 67–190

24. Karihaloo BL, Xiao QZ (2003) Linear and Nonlinear Fracture Mechanics. In: Milne I, Ritchie RO, Karihaloo BL (eds) Comprehensive structural integrity, Vol 2: Karihaloo BL, Knauss WG (eds) Fundamental theories and mechanisms of failure, Chapter 2.03. Elsevier Pergamon, UK, pp. 81–212
25. Abdalla HM, Karihaloo BL (2004) Mag Concrete Res 56: 597–604
26. Planas J, Elices M, Guinea GV, Gómez FJ, Cendón DA, Arbilla I (2003) Engrg Fract Mech 70: 1759–1776
27. Béchet E, Minnebo H, Moës N, Burgardt B (2005) Int J Numer Meth Engrg 64: 1033–1056
28. Duff IS, Reid JK (1983) ACM Trans Math Softw 9: 302–325

On the Application of Hansbo's Method for Interface Problems

Ellen Kuhl, Philippe Jäger, Julia Mergheim and Paul Steinmann

Department of Mechanical Engineering, TU Kaiserslautern, PO Box 3049,
D-67653 Kaiserslautern, Germany; ekuhl/pjaeger/mergheim/ps@rhrk.uni-kl.de

Summary. A geometrically nonlinear finite element framework for the modeling of strong discontinuities in three dimensional continua is presented. By doubling the degrees of freedom in the discontinuous elements, the algorithm allows for arbitrary discontinuities which are not a priori restricted to inter-element boundaries. On both sides of the discontinuity we apply an independent interpolation of the deformation field. Accordingly, the suggested approach relies exclusively on displacement degrees of freedom. On the discontinuity surface itself, we make use of the cohesive zone concept to account for a smooth crack opening. A three dimensional bending problem and the classical symmetric and non-symmetric peel test demonstrate the performance of the suggested method.

Key words: FE technology, failure, strong discontinuities, cohesive surfaces

1 Motivation

The modeling of strong discontinuities within the finite element setting has gained increasing attention ever since the pioneering work on extended finite elements by Belytschko and co-workers [1–6], see also Wells et al. [7–9]. The extended finite element method is extremely powerful and the underlying concept is remarkably simple: Enrichment functions are applied to interpolate a part of the deformation field which is decomposed in the standard continuous part and an additional discontinuous part. Rather then following the classical X-FEM approach and introducing jumps in the deformation field as additional unknowns, we suggest a slightly modified approach based on the recent ideas by Hansbo and Hansbo [10–12]. Their celebrated concept which essentially generalizes the ideas of Nitsche [13] was adopted by Mergheim et al. [14] and has recently been extended to finite deformations in a fully three dimensional setting, see [15].

In line with the Hansbo method, a discontinuous approximation is accomplished by doubling the degrees of freedom in the discontinuous elements. The deformation field is then interpolated independently on both sides of

Alain Combescure et al. (eds.), IUTAM Symposium on Discretization Methods for Evolving Discontinuities, 255–265.

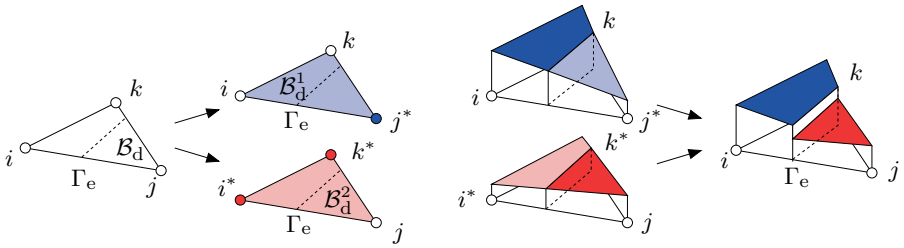


Fig. 1. Finite element discretization of strong discontinuities – Hansbo's method

the discontinuity, see figure 1. In contrast to the X-FEM, the suggested approach thus relies exclusively on deformation degrees of freedom. As such, the Hansbo method proofs particularly advantageous when treating degrees of freedom that are not additive such as directors in shell analysis as pointed out by Areias and Belytschko [6]. In contrast to the classical X-FEM, the Hansbo method is strictly local. Modifications affect only the discontinuous element itself and no additional transition elements are needed around the crack tip, as illustrated in figure 2 for the one dimensional case. Partially cracked elements, however, are more easily handled with the classical X-FEM approach by Belytschko and Black [1] and Möes et al. [2].

Once the discontinuity has been introduced, we apply the concept of cohesive surfaces which essentically dates back to the early work of Dugdale [16] and Barrenblatt [17]. Accordingly, we introduce cohesive tractions on the discontinuity surface which are constitutively related to the jump in the deformation field. This concept has been proven extremely powerful since it allows for a smooth crack opening.

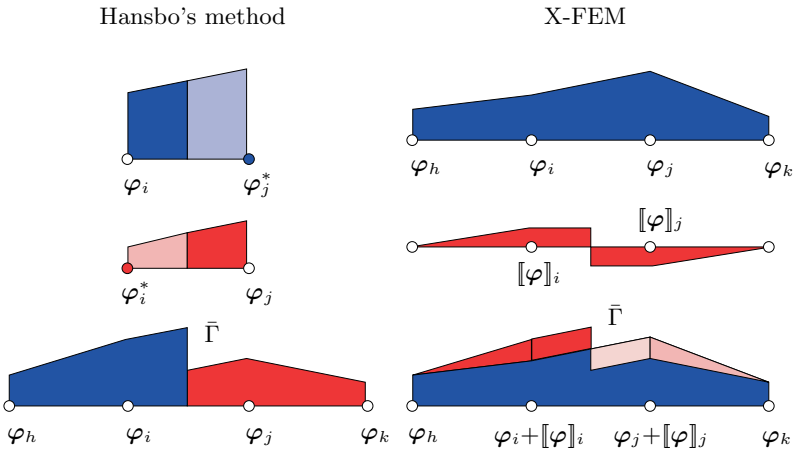


Fig. 2. Strong discontinuitites – Hansbo's method vs. classical X-FEM

Within this contribution we will focus in particular on the three dimensional treatment of strong discontinuities at large deformations. To this end, we begin by briefly summarizing the governing equations in section 2. We then address the finite element discretization in section 3. The potential of the presented scheme will be demonstrated in section 4 by means of selected examples including three point bending and the classical symmetric and non-symmetric peel test.

2 Governing Equations of Strong Discontinuities

Let $\mathcal{B}_0$ denote the material configuration occupied by of the body of interest with material placements $\boldsymbol{X}$. The corresponding spatial configuration with placements $\boldsymbol{x}$ is denoted by $\mathcal{B}$. Its boundary $\partial\mathcal{B}$ consists of disjoint parts $\partial\mathcal{B} = \partial\mathcal{B}_\varphi \cup \partial\mathcal{B}_t$ with $\partial\mathcal{B}_\varphi \cap \partial\mathcal{B}_t = \emptyset$ where either Dirichlet or Neumann boundary conditions are prescribed.

Let $\mathcal{B}_0$ be divided by a strong discontinuity Γ_0 with normal vector $\boldsymbol{N}$ introducing the two subdomains $\mathcal{B}_0^1$ and $\mathcal{B}_0^2$. By $\boldsymbol{\varphi}$ we shall denote the deformation mapping placements from the material to the spatial configuration $\boldsymbol{x} = \boldsymbol{\varphi}(\boldsymbol{X})$. Obviously, $\boldsymbol{\varphi}$ is continuous in either subdomain but exhibits a jump $[\![\boldsymbol{\varphi}]\!]$ across the discontinuity surface $\bar{\Gamma}$.

$$\boldsymbol{\varphi} = \begin{cases} \boldsymbol{\varphi}^1 \text{ in } \mathcal{B}^1 \\ \boldsymbol{\varphi}^2 \text{ in } \mathcal{B}^2 \end{cases} \qquad [\![\boldsymbol{\varphi}]\!] = \boldsymbol{\varphi}^1 - \boldsymbol{\varphi}^2 \text{ on } \bar{\Gamma} \tag{1}$$

Elements of the corresponding tangent spaces are mapped by the deformation gradient $\boldsymbol{F} = \nabla_{\boldsymbol{X}}\boldsymbol{\varphi}$ as $\mathrm{d}\boldsymbol{x} = \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{X}$. Due to the discontinuity in the deformation map $\boldsymbol{\varphi}$, the unique discontinuity surface Γ_0 is mapped onto the surfaces Γ^1 and Γ^2, see figure 3. We thus define a fictitious discontinuity surface $\bar{\Gamma}$ in the current configuration. It can be identified through the average deformation map $\bar{\boldsymbol{\varphi}} = \kappa^1\boldsymbol{\varphi}^1 + \kappa^2\boldsymbol{\varphi}^2$ and its deformation gradient $\bar{\boldsymbol{F}}$ in terms of the weighting factors κ^α which obviously have to sum up to one.

$$\boldsymbol{F} = \begin{cases} \boldsymbol{F}^1 = \nabla_{\boldsymbol{X}}\boldsymbol{\varphi}^1 \text{ in } \mathcal{B}^1 \\ \boldsymbol{F}^2 = \nabla_{\boldsymbol{X}}\boldsymbol{\varphi}^2 \text{ in } \mathcal{B}^2 \end{cases} \qquad \bar{\boldsymbol{F}} = \kappa^1\boldsymbol{F}^1 + \kappa^2\boldsymbol{F}^2 \text{ on } \bar{\Gamma} \tag{2}$$

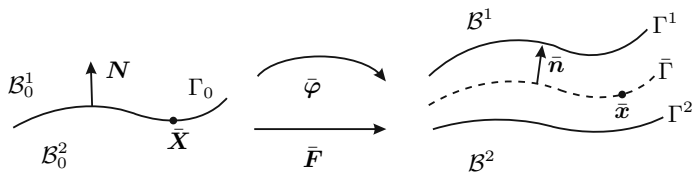


Fig. 3. Kinematics of strong discontinuities – Concept of ficititious discontinuity

For each continuous subdomain $\mathcal{B}_0^\alpha$ we thus introduce an independent field of Jacobians $J^\alpha = \det(\boldsymbol{F}^\alpha)$ and of characteristic strain measures, e.g. the field of left Cauchy Green tensors $\boldsymbol{b}^\alpha = \boldsymbol{F}^\alpha \cdot \boldsymbol{F}^{\alpha\mathrm{t}}$ with $\alpha = 1, 2$. On the discontinuity surface $\bar{\Gamma}$, the Jacobian is introduced as $\bar{J} = \det(\bar{\boldsymbol{F}})$. The set of governing equations is finally closed by setting up appropriate constitutive equations for the bulk stress $\boldsymbol{\sigma}$ and for the cohesive tractions $\bar{\boldsymbol{t}}$.

$$\boldsymbol{\sigma}(\boldsymbol{F}) = \begin{cases} \boldsymbol{\sigma}^1(\boldsymbol{F}^1) \text{ in } \mathcal{B}^1 \\ \boldsymbol{\sigma}^2(\boldsymbol{F}^2) \text{ in } \mathcal{B}^2 \end{cases} \qquad \bar{\boldsymbol{t}} = \bar{\boldsymbol{t}}_{\mathrm{n}}([\![\boldsymbol{\varphi}_{\mathrm{n}}]\!]) + \bar{\boldsymbol{t}}_{\mathrm{m}}([\![\boldsymbol{\varphi}_{\mathrm{m}}]\!]) \text{ on } \bar{\Gamma} \tag{3}$$

In what follows, we shall assume that the Cauchy stress in the bulk obeys a Neo-Hooke type constitutive law with $\boldsymbol{\sigma}^\alpha = [\,\lambda \ln(J^\alpha)\,\boldsymbol{I} - \mu\,\boldsymbol{I} + \mu\,\boldsymbol{b}^\alpha\,] / J^\alpha$ on either side $\mathcal{B}^\alpha$ of the discontinuity. The inelastic behavior is thus attributed exclusively to the discontinuity surface through the cohesive crack concept. For the sake of simplicity, we adopt an uncoupled traction separation law for which the normal traction vector $\bar{\boldsymbol{t}}_{\mathrm{n}}$ is expressed only in terms of normal jump vector $[\![\boldsymbol{\varphi}_{\mathrm{n}}]\!] = [[\![\boldsymbol{\varphi}]\!] \cdot \boldsymbol{n}]\,\boldsymbol{n}$, e.g. as $\bar{\boldsymbol{t}}_{\mathrm{n}} = f_{\mathrm{t}} \exp(-f_{\mathrm{t}} / G_{\mathrm{f}}\,[\![\boldsymbol{\varphi}]\!] \cdot \boldsymbol{n})\,\boldsymbol{n}$. Here, f_{t} denotes the tensile strength and G_{f} is the fracture energy. Similarly, the in plane traction vector $\bar{\boldsymbol{t}}_{\mathrm{m}}$ is assumed to be a function of the tangential jump $[\![\boldsymbol{\varphi}_{\mathrm{m}}]\!] = [\![\boldsymbol{\varphi}]\!] - [\![\boldsymbol{\varphi}_{\mathrm{n}}]\!]$ alone, e.g. in its most simple linear form as $\bar{\boldsymbol{t}}_{\mathrm{m}} = d\,[\![\boldsymbol{\varphi}_{\mathrm{m}}]\!]$ in terms of the shear stiffness d.

The weak formulation of the governing equations, i.e. the equilibrium equation in $\mathcal{B}_0$, the Neumann boundary conditions on $\partial\mathcal{B}_{0t}$ and the traction continuity equation along Γ_0, follows straightforwardly from the multiplication with the vector valued test function $\delta\boldsymbol{\varphi}$, the integration over the corresponding domains, the standard integration by parts and the application of the Gauss theorem. Its push forward to the spatial configuration renders the following expression

$$\int_{\mathcal{B}^1 \cup \mathcal{B}^2} \nabla_{\boldsymbol{x}} \delta\boldsymbol{\varphi} : \boldsymbol{\sigma}\,\mathrm{d}v + \int_{\bar{\Gamma}} [\![\delta\boldsymbol{\varphi}]\!] \cdot \bar{\boldsymbol{t}}\,\mathrm{d}\bar{a} - \int_{\partial\mathcal{B}_t} \delta\boldsymbol{\varphi} \cdot \boldsymbol{t}^{\mathrm{p}}\,\mathrm{d}a = 0 \tag{4}$$

in terms of the true stresses $\boldsymbol{\sigma}$, the true cohesive tractions $\bar{\boldsymbol{t}}$ and the push forward of the prescribed surface tractions $\boldsymbol{t}^{\mathrm{p}}$, see e.g. Mergheim et al. [14,15] for further details.

3 Discretization of Strong Discontinuities

In line with the Hansbo method, a discontinuous approximation is accomplished by doubling the degrees of freedom in the discontinuous elements. Each of these two independent sets of unknowns accounts for the interpolation of the deformation field on either side of the discontinuity surface. Two sets of the standard basis functions are used, one is put to zero on one side of the discontinuity while it takes its standard values on the other side and vice

versa. This leads to a formulation which allows for an arbitrarily oriented discontinuity within an element by using only displacement degrees of freedom and the standard basis functions. In the spirit of the finite element method, the domain of interest $\mathcal{B}$ is divided into n_{el} elements. For the computational tretament it proves convenient to distinguish between the standard continuous elements $\mathcal{B}_{\mathrm{c}}$, the discontinuous elements $\mathcal{B}_{\mathrm{d}}$ which are intersected by the discontinuity surface and the discontinuitiy surface $\bar{\Gamma}$ itself. In the continuous elements, the deformation map $\boldsymbol{\varphi}$ takes its classical elementwise representation. In the discontinuous elements, $\boldsymbol{\varphi}^1$ and $\boldsymbol{\varphi}^2$ are defined independently in $\mathcal{B}_{\mathrm{d}}^1$ and $\mathcal{B}_{\mathrm{d}}^2$, respectively. On the ficititious discontinuity $\bar{\Gamma}$, the jump in the deformation map $[\![\boldsymbol{\varphi}]\!]$ is described as the difference of the two continuous fields $\boldsymbol{\varphi}^1$ and $\boldsymbol{\varphi}^2$.

$$\boldsymbol{\varphi}|_{\mathcal{B}_{\mathrm{c}}} = \sum_{i=1}^{n_{\mathrm{en}}} N^i \boldsymbol{\varphi}_i \qquad \boldsymbol{\varphi}^\alpha|_{\mathcal{B}_{\mathrm{d}}^\alpha} = \sum_{i=1}^{n_{\mathrm{en}}} N^{\alpha i} \boldsymbol{\varphi}_i^\alpha \qquad [\![\boldsymbol{\varphi}]\!]|_{\bar{\Gamma}} = \sum_{i=1}^{n_{\mathrm{en}}+n_{\mathrm{en}}^*} \bar{N}^i \boldsymbol{\varphi}_i \tag{5}$$

Herein, N^i denotes the standard shape functions for tetrahedral elements, n_{en} is the number of element nodes and n_{en}^* are the nodes of the discontinuous elements which have been doubled. With $\alpha = 1, 2$ for either side of the discontinuity, $\boldsymbol{\varphi}_i^\alpha$ are the nodal values of the deformation map belonging to $\mathcal{B}_{\mathrm{d}}^\alpha$. We thus apply two copies of the standard basis functions. On the corresponding side $\mathcal{B}_{\mathrm{d}}^\alpha$ of the discontinuity, each set takes its usual values $N^{\alpha i} = N^i$ while it is set to zero as $N^{\alpha i} = 0$ on the other side. The newly introduced set $\bar{N}$ comprises the standard shape functions N^α, evaluated at $\bar{\Gamma}$ equipped with the corresponding plus or minus sign in order to represent the jump terms. The corresponding deformation gradients $\boldsymbol{F}$, $\boldsymbol{F}^\alpha$ and $\bar{\boldsymbol{F}}$ then follow naturally

$$\boldsymbol{F}|_{\mathcal{B}_{\mathrm{c}}} = \sum_{i=1}^{n_{\mathrm{en}}} \boldsymbol{\varphi}_i \otimes \nabla_{\boldsymbol{X}} N^i \quad \boldsymbol{F}^\alpha|_{\mathcal{B}_{\mathrm{d}}^\alpha} = \sum_{i=1}^{n_{\mathrm{en}}} \boldsymbol{\varphi}_i^\alpha \otimes \nabla_{\boldsymbol{X}} N^{\alpha i} \quad \bar{\boldsymbol{F}}|_{\bar{\Gamma}} = \sum_{i=1}^{n_{\mathrm{en}}+n_{\mathrm{en}}^*} \boldsymbol{\varphi}_i \otimes \kappa^i \nabla_{\boldsymbol{X}} \bar{N}^i \tag{6}$$

whereby the weighting factors κ^i can either be set to one half or be related to the corresponding element volume fractions. With these definitions at hand, the finite element discretization of equation (4) is rather straightforward. More details about the algorithmic treatment, the discrete residual equation and its consistent linearization can be found in [14, 15].

It should be pointed out that the extension from two dimensional to three dimensional crack propagation algorithms is anything but straightforward and a number of additional difficulties have to be faced when real three dimensional problems are being solved. Some of them have been addressed recently by Gasser and Holzapfel [18–20]. For example, a more cumbersome computational treatment is related to the fact that unlike two dimensional triangular elements, three dimensional tetrahedra can be intersected by the discontinuity in two different ways. If the discontinuity surface cuts three edges of the element such that the element interface becomes triangular, one tetrahedral

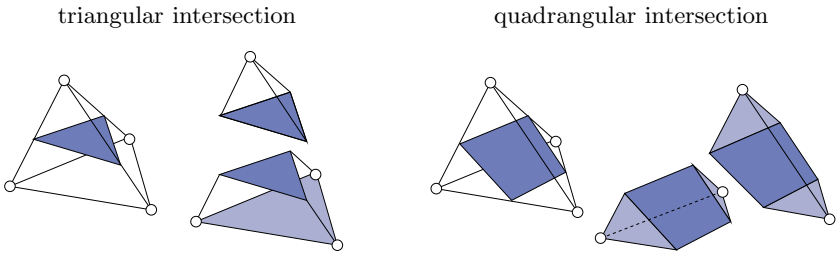


Fig. 4. Finite element discretization of strong discontinuities – Linear tetraeder

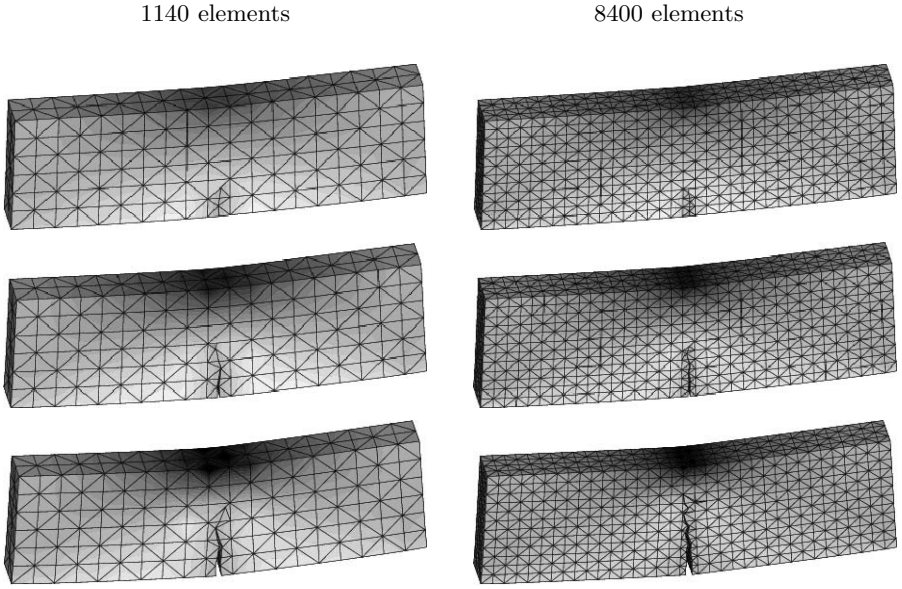


Fig. 5. Strong discontinuities – Three point bending

element and one wedge are generated, see figure 4, left. Alternatively, the discontinuity surface could intersect four edges of the element generating a quadrilateral interface as illustrated in figure 4, right. In this case, the resulting subelements are both polyhedral wedge elements.

Another yet not satisfactorily solved problem is the tracking or rather propagation of the discontinuity surface in three dimensions. Standard maximum principal stress criteria which have been applied successfully in two dimensional simulations result in non-unique crack propagation criteria in three dimensional problems. A complex global characterization of the failure surface seems to be become necessary to allow for a smooth and not yet too stiff discontinuity description, see e.g. Gasser and Holzapfel [20].

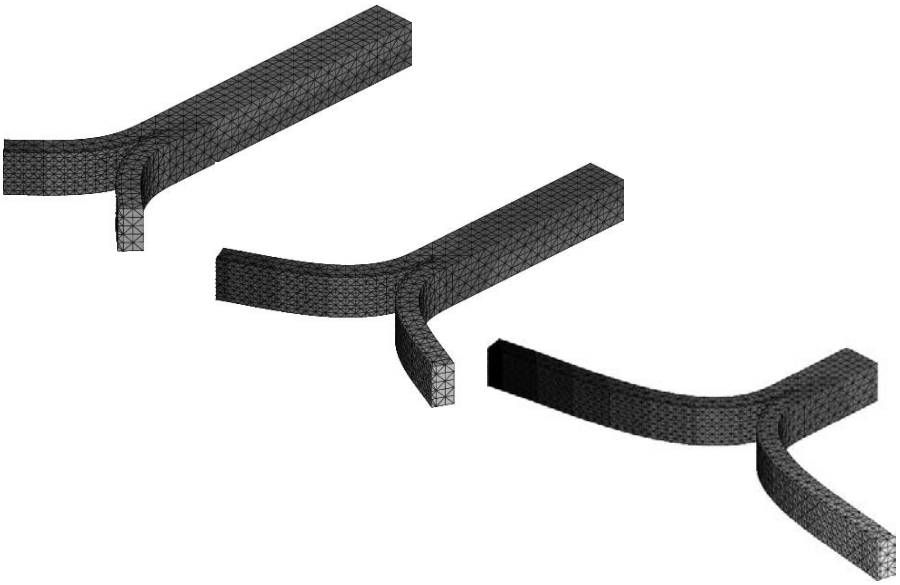


Fig. 6. Strong discontinuities – Symmetric peel test

4 Simulations of Strong Discontinuities – 3D Examples

In what follows, the potential of the suggested crack propagation algorithm is illustrated for three characteristic benchmark problems in the three dimensional setting. Figure 5 shows the simulation of the classical three point bending problem based on a finite element discretization with 1140 and 8400 linear tetrahedral elements, respectively. The beam is fixed on its two lower edges and loaded vertically on the upper mid nodes. A crack has been initiated opposite the loading. Upon a progressive increase of the load, the crack propagates smoothly from the bottom to the top as documented by the series of deformation stages depicted in figure 5. At the final stages of the loading history a clear crack opening can be observed in the tensile regime. The comparison of both crack propagation series indicates that the numerical solution is clearly mesh independent.

Next, we analyse the classical peel test which has been studied extensively in the related literature for two dimensional crack propagation problems, see e.g. Wells et al. [7]. The specimen is discretized with 2250, 3750 and 6250 linear tetrahedral elements, respectively. Figure 6 illustrates a sequence of the crack propagation history for the finest discretization. The smooth peeling of the two layers is clearly visible as the discontinuity progresses through the specimen.

Finally, we elaborate the non-symmetric peel test for which we peel off the top layer of the specimen while its bottom layer is fixed. Apart from the boundary conditions, the three analyzed discretizations are similar to the

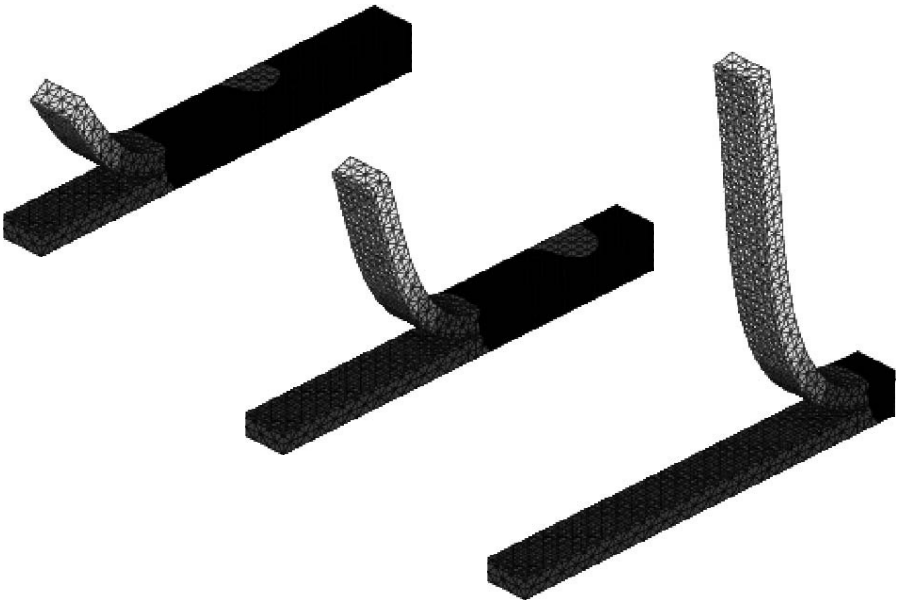


Fig. 7. Strong discontinuities – Non-symmetric peel test

previous example. The non-symmetric peel test is particularly challenging since the fictitious discontinuity surface obviously undergoes severe rotations. This benchmark problem can thus be used to elaborate the correctness of the consistent linearization of the cohesive traction term. Figure 7 shows selected stages of the simulation with the finest mesh of 6250 tetrahedral elements. Similar to the previous example, the crack opens progressively and a smooth peeling process can be observed.

Figure 8 shows the load displacement curves for both the symmetric and the non-symmetric peel test. While the 2250 element meshes seem to be too coarse to capture the failure process appropriately, the finer meshes of 3750 and 6250 elements render almost identical results. The ability to produce mesh independent results is thus a distinguishing feature or the suggested algorithm. For detailed mesh independency studies the reader is referred to Mergheim et al. [14, 15].

5 Discussion

Motivated by Hansbo's method, we presented a computational strategy for the simulation of propagating discontinuities in three dimensional continua undergoing large deformations. It is essentially characterized through discontinuous finite elements in which the deformation field is continuous on either

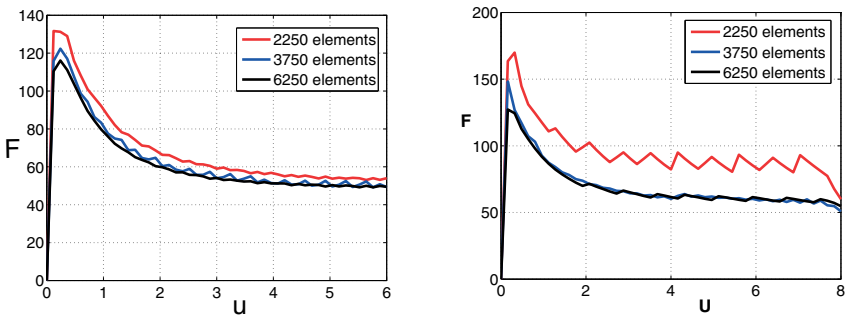


Fig. 8. Load displacement curves – Symmetric and non-symmetric peel test

side of the discontinuity surface but exhibits a jump across it. The number of degrees of freedom in these elements is doubled and the deformation is interpolated independently on both sides of the discontinuity. Smooth crack opening is allowed for through the cohesive crack concept introducing cohesive tractions which decrease exponentially upon increased crack opening.

A three point bending problem and the classical symmetric and non-symmetric peel test have been analyzed to illustrate the basic features of the suggested approach. Similar to the celebrated extended finite element method, the crack path is independent of the underlying finite element mesh. While the X-FEM requires additional transition elements, our modifications are strictly local and only affect the discontinuous elements themselves. The suggested method is thus believed to be extremely powerful in simulating propagating discontinuities not only in two but also in three dimensional continua.

Unlike for two dimensional simulations, the question of crack propagation is not yet satisfactorily solved for three dimensional problems. While nonlocal principal stress criteria are usually a good choice for two dimensional problems, these typically result in non-unique definitions for the crack propagation criterion in three dimensions. The comparison of different local and global techniques to track three dimensional crack paths appropriately is part of ongoing research.

Acknowledgements

The present research has been carried out within the *DFG Graduate School 814* sponsored through the German Science Foundation the support of which is gratefully acknowledged. Also, we would like to thank Peter Hansbo for stimulating discussions.

References

1. T. Belytschko and T. Black. Elastic crack growth in finite elements with minimal remeshing. *International Journal for Numerical Methods in Engineering*, 45:601–620, 1999.
2. N. Moës, J. Dolbow, and T. Belytschko. A finite element method for crack growth without remeshing. *International Journal for Numerical Methods in Engineering*, 46:131–150, 1999.
3. N. Sukumar, N. Möes, B. Moran, and T. Belytschko. Extended finite element method for three-dimensional crack modeling. *International Journal for Numerical Methods in Engineering*, 48(11):1549–1570, 2000.
4. N. Moës and T. Belytschko. Extended finite element method for cohesive crack growth. *Engineering Fracture Mechanics*, 69(7):813–833, 2002.
5. P. M. A. Areias, and T. Belytschko. Analysis of three-dimensional crack initiation and propagation using the extended finite element method. *International Journal for Numerical Methods in Engineering*, 63:760–788, 2005.
6. P. M. A. Areias and T. Belytschko. Letter to the Editor: A review on the article 'A finite element method for the simulation of strong and weak discontinuities in solid mechanics'. *Computer Methods in Applied Mechanics and Engineering*, 195:1275–1276, 2006.
7. G. N. Wells. Dislocation modelling of strain localization and failure. *PhD Thesis. TU Delft. The Netherlands*, 2001.
8. G. N. Wells and L. J. Sluys. A new method for modelling cohesive cracks using finite elements. *International Journal for Numerical Methods in Engineering*, 50:2667–2682, 2001.
9. G. N. Wells, L. J. Sluys and R. de Borst. A consistent geometrically non-linear approach for delamination. *International Journal for Numerical Methods in Engineering*, 54:1333–1355, 2002.
10. A. Hansbo and P. Hansbo. An unfitted finite element method, based on Nitsche's method, for elliptic interface problems. *Computer Methods in Applied Mechanics and Engineering*, 191:5537–5552, 2002.
11. A. Hansbo and P. Hansbo. A finite element method for the simulation of strong and weak discontinuities in solid mechanics. *Computer Methods in Applied Mechanics and Engineering*, 193:3523–3540, 2004.
12. P. Hansbo. Nitsche's method for interface problems in computational mechanics. *GAMM Mitteilungen*, 28:183–206, 2005.
13. J. Nitsche. Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind. *Abh. Math. Univ. Hamburg*, 36:9–15, 1970.
14. J. Mergheim, E. Kuhl, and P. Steinmann. A finite element method for the computational modelling of cohesive cracks. *International Journal for Numerical Methods in Engineering*, 63(2):276–289, 2005.
15. J. Mergheim, E. Kuhl, and P. Steinmann. Towards the algorithmic treatment of 3D strong discontinuities. *Communications Numerical Methods in Engineering*, available online first, DOI 10.1002/cnm.885.
16. D. S. Dugdale. Yielding of steel sheets containing slits. *Journal of Mechanics and Solids*, 8:100–108, 1960.
17. G. I. Barenblatt. The mathematical theory of equilibrium of cracks in brittle fracture. *Advances in Applied Mechanics*, 7:55–129, 1962.

18. T. C. Gasser and G. A. Holzapfel. Geometrically non-linear and consistently linearized embedded strong discontinuity models for 3d problems with an application to the dissection analysis of soft biological tissues. *Computer Methods in Applied Mechanics and Engineering*, 192:5059–5098, 2003.
19. T. C. Gasser and G. A. Holzapfel. Modeling 3d crack propagation in unreinforced concrete using pufem. *Computer Methods in Applied Mechanics and Engineering*, 194:2859–2896, 2005.
20. T. C. Gasser and G. A. Holzapfel. 3D Crack propagation in unreinforced concrete. A two-step algorithm for tracking 3D crack paths. *Computer Methods in Applied Mechanics and Engineering*, 195:5198–5219, 2006.

An optimal explicit time stepping scheme for cracks modeled with X-FEM

T. Menouillard[1,2], N. Moës[3], and A. Combescure[1]

[1] LaMCoS, INSA-Lyon, CNRS UMR5259
18-20 rue des Sciences, F69621 Villeurbanne Cedex, France
thomas.menouillard@insa-lyon.fr
[2] CEA Saclay, DEN / DM2S / SEMT / DYN
91191 Gif-sur-Yvette, France
[3] GEM CNRS UMR6183, Ecole Centrale Nantes
Université de Nantes, 1 rue de la Noë, 44321 Nantes, France

Summary. This paper deals with the numerical modelling of cracks in the dynamic case using X-FEM. More precisely, we are interested in explicit algorithms. We prove that by using a specific lumping technique, the critical time step is exactly the same as if no crack were present. This somewhat improves a previous result for which the critical time step was reduced by a factor of square root of 2 from the case with no crack. The new lumping technique is obtained by using a lumping strategy initially developped to handle elements containing voids. To be precise the results obtained are only valid when the crack is only modeled by the Heaviside enrichment. Note also that the resulting lumped matrix is block diagonal (blocks of size 2 by 2).

Key words: eXtended Finite Element Method, Explicit dynamics, Critical time step, Crack propagation, Lumped mass matrix.

1 Introduction

The eXtended Finite Element Method (X-FEM) allows one to introduce a crack within an existing mesh without the need to modify the mesh. Discontinuous enrichments are introduced on elements cut by the crack using the partition of unity technique [2]. The enrichment is composed of a tip enrichment [4] and a Heaviside enrichment [11] away from the crack tip. Initially developped in the 2D setting, the method was then extended to 3D in [15]. At about the same time, two level sets were introduced to conveniently store the crack as two finite element fields [16, 8]. This level set representation of cracks was also found to be extremely convenient to handle crack growth [8] (without any remeshing).

The implementation of the X-FEM requires some enhancements of the regular FEM implementation:

Alain Combescure et al. (eds.), IUTAM Symposium on Discretization Methods for Evolving Discontinuities, 267–281.

- A specific integration scheme on enriched elements must be used [11, 3];
- The number of degrees of freedom per element depends on the location of the element with respect to the crack. The use of an objected oriented language like C++ is quite convenient in dealing with the variable number of element degrees of freedom (in space and time as the crack is growing);
- Conditioning issue are raised when many layers of elements are enriched around the crack tip (a pre-conditioning technique for this event is provided in [3]).

When dynamic loadings are applied, additional issues must be addressed for the X-FEM to work properly. For transient analysis, one usually distinguishes between implicit and explicit algorithms. The standard Newmark implicit approach is known to be unconditionally stable in the finite element context for stationary cracks. As cracks are growing and remeshing is used, stability may be obtained provided that the energy release rate at the crack tip is properly computed [13].

A nice thing about the extended finite element method is that the mesh is held fixed as the crack is growing. The number of degrees of freedom is however growing during the computation. The stability of X-FEM with implicit algorithms was studied in [12] and with explicit algorithms in [10]. The latter paper introduces a special lumping technique leading to a critical time step smaller (factor $1/\sqrt{2}$) than the one in the FEM case. We shall in this paper introduce a new lumping technique allowing one to use the same critical time step as in FEM (considering only Heaviside enrichment).

The paper is organized as follows. In section 2, the critical time step is theoretically studied with two different lumping techniques for simple elements (one-dimensional, triangular and tetrahedral elements). Section 3 is dedicated to numerical experiments.

2 Explicit dynamics for X-FEM

It is well known that the critical time step of an explicit integration scheme for dynamics may be estimated by computing the minimum value of the critical time step for all elements in the mesh taken separately. Moreover, the estimate is an upper bound for the exact critical time step. This result was obtained in [6] as an extension of a theorem first given by Rayleigh.

The paragraph above remains valid if the approximation is obtained through the extended finite element method. We can thus concentrate on a single cracked element.

2.1 1D bar element

The shape functions

We start by a 1D cracked element shown in figure 1. The length of the element is L, its density ρ, its section S and its Young modulus E. The crack cuts

the element into two sub-segments of size ϵL and $(1-\epsilon)L$, respectively. Using the Heaviside enrichment [11], four approximation functions are defined over the element (figure 1): two classical and two enriched. We have assumed that the "Heaviside" type enrichment function is $+1$ on the left side of the bar and -1 on the right side. Equivalently, the four functions shown in figure 2 may be used. The latter base is in the spirit of Hansbo and Hansbo [9]. The equivalence of the basis depicted in figures 1 and 2 was proven in [1]. The basis functions are related by:

$$\begin{pmatrix} f_I \\ f_{I'} \\ f_{II} \\ f_{II'} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} f_1 \\ f_{1'} \\ f_2 \\ f_{2'} \end{pmatrix} = \mathbb{P}. \begin{pmatrix} f_1 \\ f_{1'} \\ f_2 \\ f_{2'} \end{pmatrix} \tag{1}$$

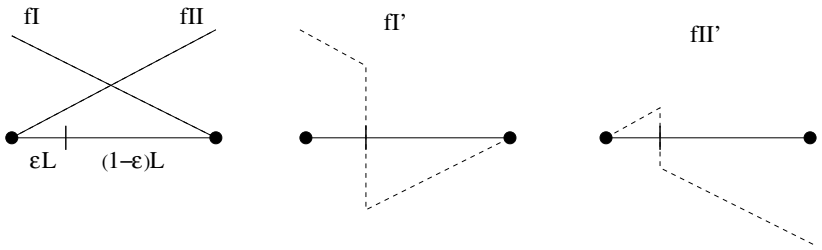


Fig. 1. Classical (f_I, f_{II}) and enriched $(f_{I'}, f_{II'})$ approximation functions on a cracked 1D element.

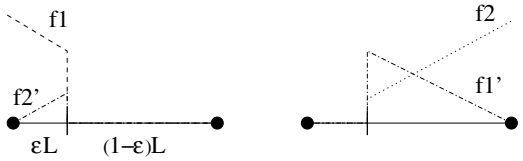


Fig. 2. Truncated function basis for a 1D cracked element (shadow node version). This basis is equivalent to the one depicted in figure 1.

Using the approximation functions depicted in figure 1, an approximated field u on the 1D element reads:

$$u = u_I f_I + u_{I'} f_{I'} + u_{II} f_{II} + u_{II'} f_{II'} \tag{2}$$

Alternatively, one can use the equivalent basis depicted in figure 2. The approximation reads with this basis:

$$u = u_1 f_1 + u_{1'} f_{1'} + u_2 f_2 + u_{2'} f_{2'} \tag{3}$$

Using the link (1), the degrees of freedom of the approximations are related by:

$$\begin{pmatrix} u_1 \\ u_{1'} \\ u_2 \\ u_{2'} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} u_I \\ u_{I'} \\ u_{II} \\ u_{II'} \end{pmatrix} = \mathbb{P}^T . \begin{pmatrix} u_I \\ u_{I'} \\ u_{II} \\ u_{II'} \end{pmatrix} \tag{4}$$

It can be noticed that the basis functions used in figure 2 are as if the two sub-segments were completely independent. The approximation on the left segment is as if the right one was void and conversely. On each sub-segment the approximation is exactly the one used to model voids with non-matching meshes as described in [7]. It is thus tempting to use the lumping strategy developped in [14] for voids leading to a critical time step independent of ϵ and of value h/c where c is the wave speed $c = \sqrt{E/\rho}$, i.e. the same critical time step as for a regular element.

The mass and stiffness matrix

Following [14], the lumped mass matrix is in our case:

$$\mathbb{M}_{1,1',2,2'} = \frac{\rho S L}{2} \begin{pmatrix} \epsilon & 0 & 0 & 0 \\ 0 & 1-\epsilon & 0 & 0 \\ 0 & 0 & 1-\epsilon & 0 \\ 0 & 0 & 0 & \epsilon \end{pmatrix} \tag{5}$$

And the stiffness matrix is written as (see [10]):

$$\mathbb{K}_{1,1',2,2'} = \frac{ES}{L} \begin{pmatrix} \epsilon & 0 & 0 & -\epsilon \\ 0 & 1-\epsilon & \epsilon-1 & 0 \\ 0 & \epsilon-1 & 1-\epsilon & 0 \\ -\epsilon & 0 & 0 & \epsilon \end{pmatrix} \tag{6}$$

The mass and stiffness matrices determine the critical time step of the element.

The critical time step

For an explicit scheme, the critical time step is computed as $2/\omega_{max}$ where ω_{max} is linked to the maximum eigenvalue of the generalized system:

$$det\left(\mathbb{K}_{1,1',2,2'} - \omega^2 \mathbb{M}_{1,1',2,2'}\right) = 0 \tag{7}$$

Thus we obtain in our case

$$\epsilon^2 (1-\epsilon)^2 \alpha^2 (\alpha-2)^2 = 0 \tag{8}$$

where

$$\alpha = \frac{\omega^2 \rho S L^2}{2ES} \tag{9}$$

So the biggest value for ω is when α is 2: $\omega_{max} = \frac{2}{L}\sqrt{\frac{E}{\rho}}$. So the critical time step is: $L\sqrt{\frac{\rho}{E}}$, which exactly corresponds to the critical time step of the finite element problem.

The results obtained here seem really interesting: the critical time step of a one-dimensional enriched element is exactly the same as the one of a standard element. Moreover this result does not depend on the shape functions basis used: $(1, 1', 2, 2')$ or (I, I', II, II'). The proof is

$$\begin{aligned} det(\mathbb{K}_{1,1',2,2'} - \omega^2 \mathbb{M}_{1,1',2,2'}) &= det(\mathbb{P}^T \mathbb{K}_{I,I',II,II'} \mathbb{P} - \omega^2 \mathbb{P}^T \mathbb{M}_{I,I',II,II'} \mathbb{P}) \\ &= det(\mathbb{P}^T (\mathbb{K}_{I,I',II,II'} - \omega^2 \mathbb{M}_{I,I',II,II'}) \mathbb{P}) \\ &= det(\mathbb{K}_{I,I',II,II'} - \omega^2 \mathbb{M}_{I,I',II,II'}) \end{aligned} \tag{10}$$

where matrix $\mathbb{P}$ is defined in equation (11). So going back to the X-FEM basis functions, we apply to the mass matrix above the transformation

$$\mathbb{M}_{I,I',II,II'} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix}^T . \mathbb{M}_{1,1',2,2'} . \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \tag{11}$$

$$= \frac{\rho S L}{2} \begin{pmatrix} 1 & 2\epsilon - 1 & 0 & 0 \\ 2\epsilon - 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2\epsilon - 1 \\ 0 & 0 & 2\epsilon - 1 & 1 \end{pmatrix} \tag{12}$$

We did obtain a block diagonal mass matrix. The blocks being of size 2.

Finally, had we decided that node 1 was on the negative side of the "Heaviside", the results would have been different. The results would have been the one obtained by replacing ϵ by $1-\epsilon$ in (12). The result (12) is thus valid whatever the choice of sign for the Heaviside, provided that ϵ is always defined as the matter fraction on the positive Heaviside side.

Kinetic energy conservation

In this part we check that the lumping techniques conserve the kinetic energy for 2 rigid modes.

In the classical finite element setting, the sum of the entries in a mass matrix is always equal to the total mass. The reason for this is that if all degrees of freedom are set to one, we obtain a rigid translatio mode. In the X-FEM case, the sum of the entries of the matrix is not equal to the total mass because setting to 1 all degrees does not give a rigid mode. However if we sum the entries corresponding to a rigid mode (firt and third line and column) we recover the total mass. Let us consider a one-dimensional element with two nodes. Each node has ordinary degrees of freedom corresponding to the shape functions N_1 and N_2 and additional degrees of freedom corresponding

to the enriched function H. The fraction ratio of this element is $\epsilon \in [0, 1]$. First we consider a rigid motion at speed V (as if there was no crack). The exact kinetic energy is:

$$E_c = \frac{1}{2} m V^2 \tag{13}$$

where $m = \rho S L$ is the mass of the element, S its section, L its length and ρ its density. And the discretized corresponding to this motion is described by the vector (in the Hansbo basis $(1, 1', 2, 2')$ as shown in figure 2):

$$[U] = \begin{bmatrix} V & V & V & V \end{bmatrix}^T \tag{14}$$

So the discretized kinetic energy reads:

$$\begin{aligned} E_{dis} &= \frac{1}{2}[U].\mathbb{M}_{1,1',2,2'}.[U] = \frac{1}{2}\frac{m}{2}\left(\epsilon + 1 - \epsilon + 1 - \epsilon + \epsilon\right) V^2 \\ &= \frac{1}{2} m V^2 = E_c \end{aligned} \tag{15}$$

Let us now consider the rigid motions of a cut element whose parts move away at speed V. The discretized speed vector is:

$$[U] = \begin{bmatrix} V & -V & -V & V \end{bmatrix}^T \tag{16}$$

And the discretized kinetic energy can be written as:

$$\begin{aligned} E_{dis} &= \frac{1}{2}[U].\mathbb{M}_{1,1',2,2'}.[U] = \frac{1}{2}\frac{m}{2}\left(\epsilon V^2 + (1-\epsilon)(-V)^2 + (1-\epsilon)(-V)^2 + \epsilon V^2\right) \\ &= \frac{1}{2} m V^2 = E_c \end{aligned} \tag{17}$$

The other lumped mass matrix has already been verified as well in [10].

Let us now consider a more complex case. The figure 3 presents a bar composed of three elements; the middle one is enriched. Each element has a length L, its Young modulus is E, its section S and its density ρ. The position of the discontinuity in the middle element is at $(1+\epsilon)L$ far from the left node of the 3-elements structure. The length of the bar is $l = 3L$, and the mass is $m = \rho S l$. Here we check the conservation of kinetic energy for the bar. The figure 4 presents the X-FEM shape functions: 4 standard and 2 enriched functions.

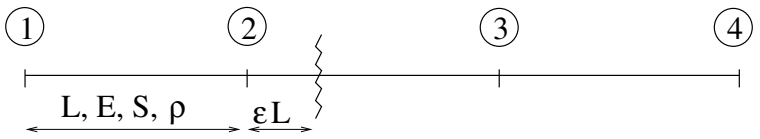


Fig. 3. A 3-elements structure cut by a crack.

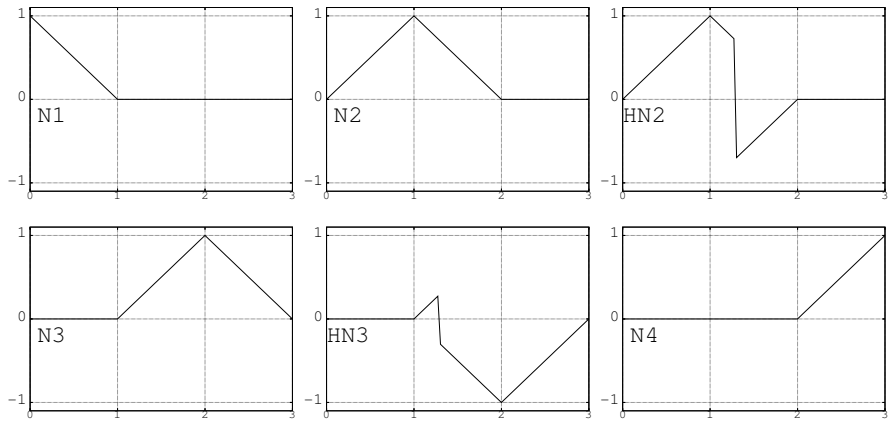


Fig. 4. Standard shape functions for the 3 elements mesh: N_1, N_2, HN_2, N_3, HN_3 and N_4.

The discrete displacement space is of size 6:

$$[U]_{diag} = \begin{bmatrix} u_1 \; u_2 \; a_2 \; u_3 \; a_3 \; u_4 \end{bmatrix}^T \tag{18}$$

and the approximate displacement is

$$\begin{aligned} U(M,t) = u_1.N_1(M) + u_2.N_2(M) + a_2.H(M).N_2(M) + u_3.N_3(M) \\ +a_3.H(M).N_3(M) + u_4.N_4(M) \end{aligned} \tag{19}$$

The diagonal mass matrix [10] for the whole structure is the following: (with $l = 3L$)

$$\mathbb{M}_{diag} = \frac{\rho SL}{2} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{20}$$

And the block diagonal mass matrix [14] is:

$$\mathbb{M}_{block-diag} = \frac{\rho SL}{2} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 2\epsilon - 1 & 0 & 0 & 0 \\ 0 & 2\epsilon - 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2\epsilon - 1 & 0 \\ 0 & 0 & 0 & 2\epsilon - 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{21}$$

$\mathbb{M}_{diag}$ is the diagonal mass matrix obtained with the technique developped in [10], and $\mathbb{M}_{block-diag}$ the block diagonal mass matrix obtained in [14]. As in the last paragraph let us consider two rigid motions.

Motion 1: rigid body at speed V. The speed vector is:

$$[v_1] = \begin{bmatrix} V & V & 0 & V & 0 & V \end{bmatrix}^T \tag{22}$$

The discretized energy is evaluated by $\frac{1}{2}[v_1]^T.\mathbb{M}.[v_1]$

$$Ec^{diag}_{motion-1} = \frac{1}{2}[v_1]^T.\mathbb{M}_{diag}.[v_1] = \frac{1}{2}\rho SlV^2 = Ec^{exact}_{motion-1} \tag{23}$$

$$Ec^{block-diag}_{motion-1} = \frac{1}{2}[v_1]^T.\mathbb{M}_{block-diag}.[v_1] = \frac{1}{2}\rho SlV^2 = Ec^{exact}_{motion-1} \tag{24}$$

Motion 2: rigid body left part at speed $-V$, and right part at speed V. So the speed vector is written as:

$$[v_2] = \begin{bmatrix} -V & 0 & -V & 0 & -V & V \end{bmatrix}^T \tag{25}$$

The discretized energy is evaluated by: $\frac{1}{2}[v_2]^T.\mathbb{M}.[v_2]$ for both mass matrix:

$$Ec^{diag}_{motion-2} = \frac{1}{2}[v_2]^T.\mathbb{M}_{diag}.[v_2] = \frac{1}{2}\rho SlV^2 = Ec^{exact}_{motion-2} \tag{26}$$

$$Ec^{block-diag}_{motion-2} = \frac{1}{2}[v_2]^T.\mathbb{M}_{block-diag}.[v_2] = \frac{1}{2}\rho SlV^2 = Ec^{exact}_{motion-2} \tag{27}$$

We have to notice that the two lumped mass matrix allow to conserve the kinetic energy for both rigid body motions.

Conclusion

To conclude on the one-dimensional case, the lumping technique developped in [14] allows one to obtain the same critical time step for an enriched element than for a standard element. This result improves the technique explained in [10] which was only $\sqrt{2}$ smaller than standard one. However the cost here to obtain a better result is in the mass matrix which depends on the fraction ratio, whereas mass matrix in [10] was simply diagonal constant.

In addition, we showed that these two lumping techniques allow to conserve kinetic energy for rigid motions.

Now the use of each technique can be motivated by different reasons: the use of block diagonal matrix will be preferable except for pure explicit codes which does not own matrix (as LS DYNA, RADIOSS, EUROPLEXUS). For these codes, a constant diagonal mass matrix will be enough unless they were able to store 2 by 2 matrices for each enriched node.

2.2 2D element: Triangular element

We now move to 2D elements, we shall see that the mass matrix keeps exactly the same topology as in (12). Consider a triangular element depicted in figure

5. As in the 1D case, we use roman indices to denote the classical shape functions and the roman prime indices to denote the enriched shape function. The arabic indices denote the truncated shape functions. For instance, the function f_1 is equal to the classical shape function f_I on the sub-triangle containing node 1 and is zero on the other sub-triangle. Finally the prime arabic indices indicate the "complementary" truncated functions. For instance, $f_{1'}$ is such that $f_1 + f_{1'} = f_I$.
The truncated and enriched basis functions are related by

$$f_I = f_1 + f_{1'} \tag{28}$$
$$f_{I'} = H(x_I)\,(f_1 - f_{1'}) \tag{29}$$
$$f_{II} = f_2 + f_{2'} \tag{30}$$
$$f_{II'} = H(x_{II})\,(f_2 - f_{2'}) \tag{31}$$
$$f_{III} = f_3 + f_{3'} \tag{32}$$
$$f_{III'} = H(x_{III})\,(f_3 - f_{3'}) \tag{33}$$

where $H(.)$ indicates the sign of the Heaviside at the corresponding node. The mass matrix in the truncated base reads

$$\mathbb{M}_{1,1',2,2',3,3'} = \frac{\rho S}{3}\begin{pmatrix} \epsilon & 0 & 0 & 0 & 0 & 0 \\ 0 & 1-\epsilon & 0 & 0 & 0 & 0 \\ 0 & 0 & 1-\epsilon & 0 & 0 & 0 \\ 0 & 0 & 0 & \epsilon & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-\epsilon & 0 \\ 0 & 0 & 0 & 0 & 0 & \epsilon \end{pmatrix} \tag{34}$$

Applying the proper transformation, based on (28-33), we obtain the following mass matrix for the enriched basis:

$$\mathbb{M}_{I,I',II,II',III,III'} = \frac{\rho S}{3}\begin{pmatrix} 1 & 2\epsilon-1 & 0 & 0 & 0 & 0 \\ 2\epsilon-1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2\epsilon-1 & 0 & 0 \\ 0 & 0 & 2\epsilon-1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2\epsilon-1 \\ 0 & 0 & 0 & 0 & 2\epsilon-1 & 1 \end{pmatrix} \tag{35}$$

We have assumed that the material fraction ϵ is computed on the positive side of the Heaviside. Otherwise, ϵ must be replaced by $1 - \epsilon$. The obtained mass matrix is basically the same as in the 1D case. For 3D elements, the 2 x 2 block diagonal matrix will not change except for the factor in front of the matrix which will be the element volume divided by the number of nodes of the element.

The results for a triangular element is presented in the figure 6. The figure 6 presents the critical time step normalized by the Finite Element critical time step obtained with a lumped mass matrix. The two results are for the

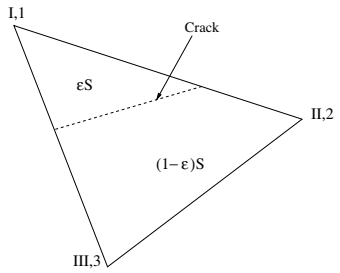


Fig. 5. A triangular element of area S cut by a crack. The fraction ratio is denoted ϵ.

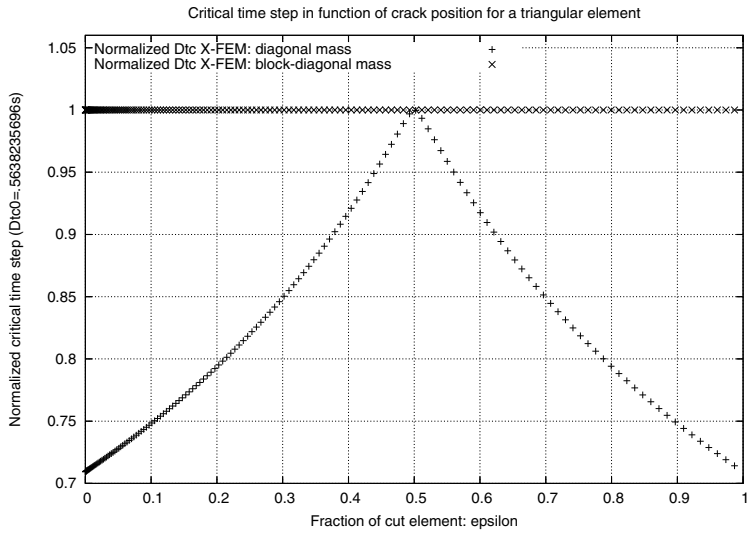


Fig. 6. Normalized critical time step for a triangular element in function of the position of the crack as explained in Figure 5.

two different lumping techniques which use a diagonal mass matrix [10], and a block diagonal mass matrix [14], respectively. It underlines the fact that the critical time step of the enriched triangular element is the same than the one of the triangular finite element, whatever the position of the crack in the element. Hence the result is similar to the one-dimensional case. Moreover the figure 6 shows that the critical time step for the diagonal case is quite smaller than for the block-diagonal case. Finally, note that the block-diagonal 2×2 matrix is not the same for all the enriched nodes. It depends on the crack path.

To conclude, using a block diagonal mass matrix, one obtains the same critical time step than in the Finite Element case, so around twice the value of the critical time step obtained using constant diagonal mass matrix. The

block diagonal mass matrix is composed of 2 by 2 matrices each of which depends on the fraction ratios ϵ of the elements around the considered nodes.

2.3 3D element: Tetrahedral element

The figure 7 presents a tetrahedral element cut by a crack. The figure 8

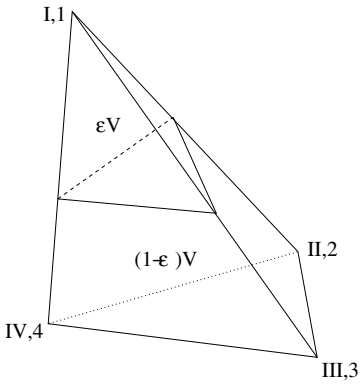


Fig. 7. Three dimensional tetrahedral element of volume V cut by a crack. ϵ is the fraction ratio on the positive side of the Heaviside function.

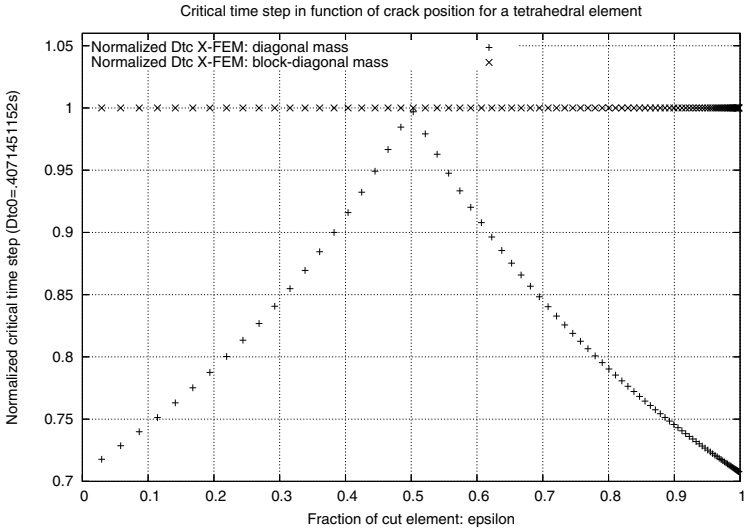


Fig. 8. Normalized critical time step for a tetrahedral element in function of the position of the crack as explained in Figure 7.

presents the critical time step normalized by the Finite Element critical time step obtained with a lumped mass matrix. The two results are for the two different lumping techniques which use a diagonal mass matrix [10], and a block diagonal mass matrix [14], respectively.

For this 3D element, the critical time step obtained with a block diagonal mass matrix is exactly the same as the one of the non-enriched Finite Element. So the results for the three dimensions studied in this paper are similar: one-dimensional linear element, triangular element and tetrahedral element allows one to obtain a X-FEM critical time step equal to the FEM critical time step.

2.4 Results analysis

In this section, we take a deeper look at the reasons why we can obtain the same critical time step than for the Finite Element case. Using the Hansbo basis for an enriched element, we have for a 1D element (as it is shown on figure 2)

$$\begin{aligned}
\frac{1}{L}\int_0^L \frac{\partial f_1(x)}{\partial x}dx &= \int_0^\epsilon dX = \epsilon \\
\frac{1}{L}\int_0^L \frac{\partial f_{1'}(x)}{\partial x}dx &= \int_0^\epsilon dX = \epsilon \\
\frac{1}{L}\int_0^L \frac{\partial f_2(x)}{\partial x}dx &= \int_\epsilon^1 dX = 1-\epsilon \\
\frac{1}{L}\int_0^L \frac{\partial f_{2'}(x)}{\partial x}dx &= \int_\epsilon^1 dX = 1-\epsilon
\end{aligned} \tag{36}$$

Thus the X-FEM stiffness matrix is proportional to the FEM stiffness matrix on each side of the discontinuity.

$$\mathbb{K}_{1,2',1',2} = \frac{ES}{L}\begin{pmatrix} \epsilon & -\epsilon & 0 & 0 \\ -\epsilon & \epsilon & 0 & 0 \\ 0 & 0 & 1-\epsilon & \epsilon-1 \\ 0 & 0 & \epsilon-1 & 1-\epsilon \end{pmatrix} = \begin{pmatrix} \epsilon\mathbb{K}_{FEM} & \mathbb{O} \\ \mathbb{O} & (1-\epsilon)\mathbb{K}_{FEM} \end{pmatrix} \tag{37}$$

where $\mathbb{K}_{FEM}$ is described in the appendix. The same may be said for the lumped mass matrix written in equation (5). The fact that both stiffness and mass matrices are proportional to the FEM on each side of the discontinuity explains why the X-FEM critical time step is the same as the FEM one.

For the three simple elements we studied, the standard shape functions were linear in space, and written as the following functions Φ_k (where $k \in \{1...n_{nodes}\}$, n_{nodes} is the number of nodes in the element):

$$\Phi_k(x,y,z) = I_0 + I_1.x + I_2.y + I_3.z \tag{38}$$

So the gradient of the shape functions are constant as it was explained in equation (36). This is why the X-FEM stiffness matrix is linked to the Finite Element one.

Note that this is not the case for a quadrangular element because the shape functions are not linear:

$$\Phi_k(x, y) = I_0 + I_1.x + I_2.y + I_3.x.y \tag{39}$$

We can sum up the results in table 1 which presents the normalized critical time step for the 3 studied elements: one-dimensional linear, triangular and tetrahedral elements, with the two lumping techniques.

Element	Δt_c^{FEM}		Δt_c^{XFEM} normalized by $\Delta t_c^{FEMlump}$	
	M standard	**M** lumped	**M** block-diagonal	**M** diagonal
1D	0.577	1	1	$> 0,707$
Triangular	$0,2819$	$0,5638$	1	$> 0,707$
Tetrahedral	$0,1919$	$0,4071$	1	$> 0,707$

Table 1. Table presenting critical time steps for different elements: standard and enriched (Young's modulus $E = 1$, length $L = 1$, density $\rho = 1$, Poisson's ratio $\nu = 0,3$).

3 Numerical applications

The compact compression specimen (CCS) problem is described schematically in figure 9. The material properties are those of PMMA: $E = 5.76\ GPa$, $\nu = 0.42$, $\rho = 1,180 kgm^{-3}$. The force $P_1(t)$ is due to an impact at velocity $V_0 = 20ms^{-1}$. The CCS is assumed to be linear elastic. Although the CCS geometry is symmetrical, the deformation and, therefore, the crack's path is not. This is due to the non-symmetric loading and boundary conditions.

We carry the computations with the two methods: first using a block diagonal mass matrix and a time step close to the corresponding critical time step of this method, and second the technique using a diagonal mass matrix with its corresponding critical time step. For both the results agree with the experiments [17, 18]: crack path and velocity of crack tip. Figure 10 shows the final deformed shape.

The technique with the diagonal mass matrix is interesting because it does not require inversion of 2 by 2 matrices at the enriched nodes. On the other hand, the technique with the diagonal mass leads to a time step quite smaller. The time step of the block diagonal method (bigger than the other one) largely allows to compensate the time used for the inversion of the block diagonal mass matrix. This fact is even more obvious as the non linearities appear as more time is spend in the computation of the internal forces.

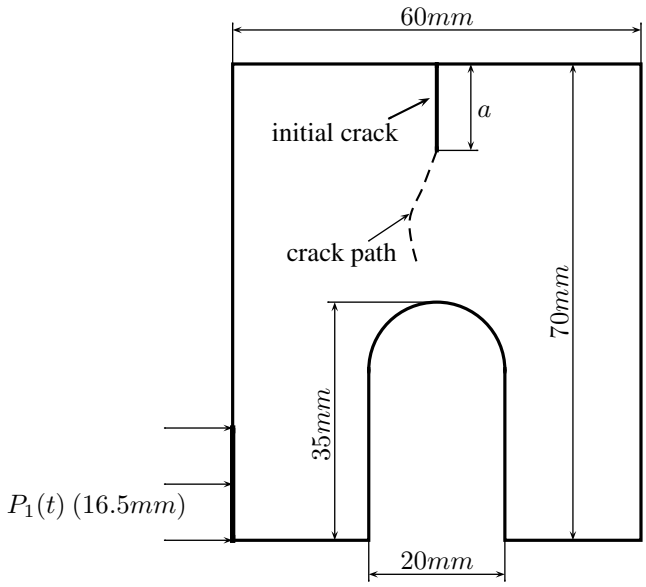


Fig. 9. Modeling of the CCS: boundary conditions and geometry (specimen thickness: 16.5mm)

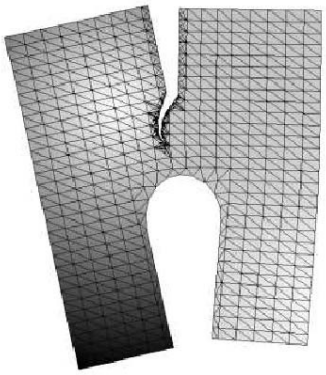


Fig. 10. Deformed mesh (with triangular elements).

4 Conclusion

In this paper, we introduced a new lumping technique for the mass matrices of meshes enriched with Heaviside functions with the X-FEM. This lumping technique yields the same critical time step at the element level than the one for non enriched elements. The lumped mass matrix is not strictly diagonal but rather block diagonal. A 2 by 2 matrix needs to be stored at each enriched node. This additional storage provides a better critical time step than the

one obtained using a true diagonal mass matrix. Numerical experiments did demonstrate the robustness and stability of the approach.

Appendix: One-dimensional Finite Element problem

For the 1D Finite Element problem, the consistent mass matrix, the lumped mass matrix and the stiffness matrix are:

$$\mathbb{M}_{FEM} = \frac{\rho SL}{6}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \qquad \mathbb{M}_{FEM}^{lumped} = \frac{\rho SL}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \mathbb{K}_{FEM} = \frac{ES}{L}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

So the corresponding stable time step is: $\Delta t_c^{lumped} = L\sqrt{\frac{\rho}{E}} = \sqrt{3}\Delta t_c^{consistent}$

References

1. Areias P., Belytschko T. (2006) Computer Methods In Applied Mechanics and Engineering 195:1275-1276
2. Babuka I., Melenk I. (1997) International Journal for Numerical Methods in Engineering 40/4:727-758
3. Béchet E., Minnebo H., Moës N., Burgardt B. (2005) International Journal for Numerical Methods in Engineering 64:1033-1056
4. Belytschko T., Black T. (1999) International Journal for Numerical Methods in Engineering 45/5:601-620
5. Belytschko T., Chen H., Xu J. X., Zi G. (2003) International Journal For Numerical Methods in Engineering 58:1873-1905
6. Belytschko T., Smolinski P., Liu W. (1985) Computer Methods In Applied Mechanics and Engineering 49:281-297
7. Daux C., Moës N., Dolbow J., Sukumar N., Belytschko T. (2000) International Journal for Numerical Methods in Engineering 48:1741-1760
8. Gravouil A., Moës N., Belytschko T. (2002) International Journal for Numerical Methods in Engineering 53:2569-2586
9. Hansbo A., Hansbo P. (2004) Computer Methods In Applied Mechanics and Engineering 193:3523-3540
10. Menouillard T., Réthoré J., Combescure A., Bung H. (2006) International Journal for Numerical Methods in Engineering 68:911–939
11. Moës N., Dolbow J., Belytschko T. (1999) International Journal for Numerical Methods in Engineering 46:131-150
12. Réthoré J., Gravouil A., Combescure A. (2005) International Journal for Numerical Methods in Engineering 63/5:631–659
13. Réthoré J., Gravouil A., Combescure A. (2004) Computer Methods in Applied Mechanics and Engineering 193:4493–4510
14. Rozycki P., Moës N., Béchet E., Dubois C. (2006) Computer Methods in Applied Mechanics and Engineering, Special Issue, to appear.
15. Sukumar N., Moës N., Moran B., Belyschko T. (2000) International Journal for Numerical Methods in Engineering 48:1549–1570
16. Stolarska M., Chopp D. L., Moës N., Belytschko T. (2001)International Journal for Numerical Methods in Engineering 51:943–960
17. Rittel D., Maigre H. (1996) Mechanics of Materials 23:229–239
18. Rittel D., Maigre H. (1996) Mechanics Research Communications 23:475–481

Variational Extended Finite Element Model for Cohesive Cracks: Influence of Integration and Interface Law

Günther Meschke[1], Peter Dumstorff[2], and Wagner Fleming[1]

[1] Institute for Structural Mechanics, Ruhr-University Bochum, Universitätsstrasse 150, 44780 Bochum, Germany
guenther.meschke@rub.de
[2] Siemens Power Generation, 45478 Mülheim an der Ruhr, Germany

Summary. According to a recently proposed variational formulation of the Extended Finite Element Method for cohesive crack propagation analyses [1] the length and the direction of new crack segments are determined on a global level from minimizing the total energy of the system. The focus of this paper is laid on the influence of the numerical integration and of the cohesive interface law on the energy distribution and, consequently, on the predicted crack trajectory. These influences are investigated by means of crack propagation analyses in plane structures made of quasi-brittle materials.

Key words: Extended Finite Element Method, variational formulation, crack propagation, fracture, cohesive cracks, quasi-brittle materials, crack interface

1 Introduction

Reliable prognoses of serviceability and of ultimate states of structures made of quasi-brittle materials such as concrete require robust and reliable computational models for the opening and propagation of cracks. Numerous investigations by different authors have shown that the Extended Finite Element Method (X-FEM) is capable of simulating cracking and crack propagation independently of the discretisation of the finite element mesh [2–4]. The X-FEM was first introduced in the context of linear elastic fracture mechanics [2] and has been extended for the modelling of cohesive cracks [3, 4].

The analysis of crack propagation crucially depends on the crack growth criterion. In most of the existing implementations of X-FEM-based models for cohesive cracks, the opening of new cracks is based upon a local stress-based criterion, such as the principal stress, evaluated at the crack tip, while the direction of crack propagation is determined according to the principal axes of stresses evaluated in the vicinity of the crack tip by means of a non-

Alain Combescure et al. (eds.), IUTAM Symposium on Discretization Methods for Evolving Discontinuities, 283–301.

local averaging procedure [4–6]. An alternative to this local approach was proposed by OLIVER and co-workers [7,8] with the so-called ,,Global Tracking Algorithm".

In this paper a variational formulation of the Extended Finite Element Method for propagation of cohesive cracks in quasi-brittle solids recently proposed in [1] is investigated numerically. According to this model crack propagation (crack direction as well as the length of new crack segments) is determined on a global level from minimizing the total energy of the system. Since in this model the energy distribution controls the predicted crack trajectory, main emphasis of this paper is laid on the influence of the numerical integration (Section 4) and of the interface law (Section 5) on the energy distribution. Starting from a variational formulation in terms of the displacement field, the directions and the length of new crack segments, a coupled format of the tangential algebraic equation system analogous to multifield problems is solved simultaneously by means of the NEWTON iteration scheme. The idea of revisiting brittle fracture as an energy minimization problem has been formulated previously (restricted to cohesionless cracks) in a general variational format in [9]. A comparative assessment of the aforementioned local and global approaches including the energy-based Extended Finite Element model is provided in [10].

In the model investigated in this paper, cohesive cracks may continuously penetrate through elements. Hence, besides enriching elements which are fully crossed by a crack, enrichments of the displacement field in the vicinity of the crack tip (which do not lead to singular stress fields) are included.

2 Variational Formulation of Cracked Bodies

Assuming quasi-static and isothermal conditions, a body whose domain Ω contains an existing curved cohesive crack and a new crack segment, creating a new crack surface A_c, which propagates with a kinking angle θ_c relative to the existing segment is considered. The total area of the current configuration of the crack is denoted as Γ_c.

The total energy of this body at a certain loading stage t consists of the internal energy $\mathcal{U}$, the external work $\mathcal{W}_b$ and $\mathcal{W}_t$ of the body forces $\boldsymbol{b}$ and the surface tractions $\boldsymbol{t}^\star$, respectively, and the surface energy of the crack $\mathcal{W}_c$:

$$\Psi(\boldsymbol{u}, A_c, \theta_c) = \mathcal{U} - \mathcal{W}_b - \mathcal{W}_t + \mathcal{W}_c, \tag{1}$$

with

$$\begin{aligned} \mathcal{U} &= \int_{\Omega \setminus \Gamma_c} \int_{\mathbf{0}}^{\boldsymbol{\varepsilon}_t} \boldsymbol{\sigma}(\boldsymbol{\varepsilon}) \mathrm{d}\boldsymbol{\varepsilon}\, \mathrm{d}V, \\ \mathcal{W}_b &= \int_{\Omega} \int_{\mathbf{0}}^{\mathbf{u}_t} \boldsymbol{b} \mathrm{d}\boldsymbol{u}\, \mathrm{d}V, \;\; \mathcal{W}_t = \int_{\Gamma_\sigma} \int_{\mathbf{0}}^{\mathbf{u}_t} \boldsymbol{t}^\star \, \mathrm{d}\boldsymbol{u}\, \mathrm{d}A. \end{aligned} \tag{2}$$

The total energy Ψ depends on the displacement field $\boldsymbol{u}$, the crack angle θ_c and the surface A_c of the created crack segment.

The displacement field of a body containing a crack can be decomposed into a continuous part $\bar{\boldsymbol{u}}$ and a discontinuous part $\check{\boldsymbol{u}}$

$$\boldsymbol{u}(\boldsymbol{x}) = \bar{\boldsymbol{u}}(\boldsymbol{x}) + \check{\boldsymbol{u}}(\boldsymbol{x}) \quad \forall\, \boldsymbol{x} \in \Omega \quad \text{with} \quad \check{\boldsymbol{u}}(\boldsymbol{x}) = S_c(\boldsymbol{x})\, \hat{\boldsymbol{u}}(\boldsymbol{x}), \tag{3}$$

where $\bar{\boldsymbol{u}}$ and $\hat{\boldsymbol{u}}$ are continuous functions in the domain Ω and S_c is the Sign function. Inserting the displacement field (3) into $(2)_1$ and assuming linear elastic behavior of the uncracked material yields:

$$\begin{aligned} \mathcal{U} &= \tfrac{1}{2} \int_{\Omega \setminus \Gamma_c} \boldsymbol{\varepsilon} : \mathbb{C} : \boldsymbol{\varepsilon}\, \mathrm{d}V \\ &= \tfrac{1}{2} \int_{\Omega \setminus \Gamma_c} (\boldsymbol{\nabla}^s \bar{\boldsymbol{u}} : \mathbb{C} : \boldsymbol{\nabla}^s \bar{\boldsymbol{u}} + \boldsymbol{\nabla}^s \hat{\boldsymbol{u}} : \mathbb{C} : \boldsymbol{\nabla}^s \hat{\boldsymbol{u}} + 2 S_c \boldsymbol{\nabla}^s \bar{\boldsymbol{u}} : \mathbb{C} : \boldsymbol{\nabla}^s \hat{\boldsymbol{u}})\, \mathrm{d}V. \end{aligned} \tag{4}$$

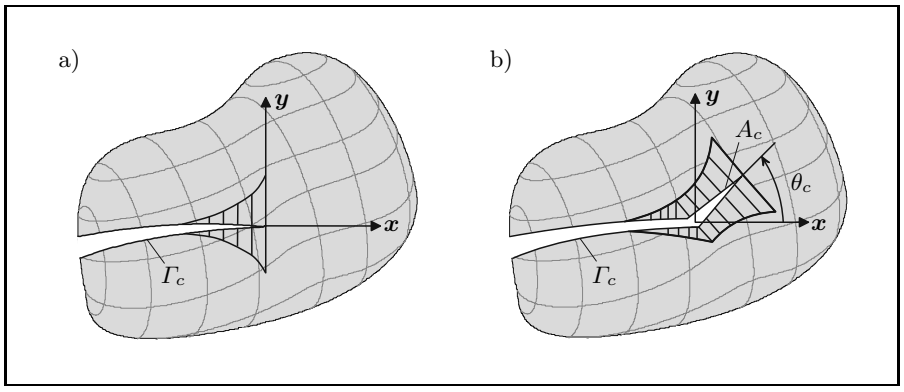


Fig. 1. Schematic illustration of crack propagation in a cracked body

For cohesive cracks, the surface energy $\mathcal{W}_c$ is given as

$$\mathcal{W}_c = \int_{\Gamma_c} \int_0^{[\![\mathbf{u}]\!]_t} \boldsymbol{t}([\![\boldsymbol{u}]\!])\, \mathrm{d}[\![\boldsymbol{u}]\!]\, \mathrm{d}A, \tag{5}$$

with the separation-dependent residual traction vector $\boldsymbol{t}([\![\boldsymbol{u}]\!])$ acting along the process zone of the crack.

Among all possible deformed configurations of the body containing one (or more) crack(s) extended by one (or more) new crack segment(s), the actual one, associated with safe equilibrium of the cracked body will lead to a minimum of $\Psi(\boldsymbol{u}, \theta_c, A_c)$ [11]. This is equivalent to the stationarity condition:

$$\delta\Psi(\boldsymbol{u},\theta_c,A_c)=\frac{\partial\Psi}{\partial\boldsymbol{u}}\delta\boldsymbol{u}+\frac{\partial\Psi}{\partial\theta_c}\delta\theta_c+\frac{\partial\Psi}{\partial A_c}\delta A_c=0. \tag{6}$$

In the following the consideration of body forces is neglected and it is assumed that the term $\mathcal{W}_t$ which corresponds to the applied tractions $\boldsymbol{t}^\star$ is independent of θ_c and A_c. Inserting Equation (1) and (2) into (6) leads to the conditions:

$$\begin{aligned}\int_{\Omega\setminus\Gamma_c}\boldsymbol{\sigma}:\delta\boldsymbol{\varepsilon}\,\mathrm{d}V-\int_{\Gamma_\sigma}\boldsymbol{t}^\star\delta\boldsymbol{u}\,\mathrm{d}A+\int_{\Gamma_c}\boldsymbol{t}\frac{\partial[\![\boldsymbol{u}]\!]}{\partial\boldsymbol{u}}\delta\boldsymbol{u}\,\mathrm{d}A&=0,\\ \int_{\Omega\setminus\Gamma_c}\boldsymbol{\sigma}:\frac{\partial\boldsymbol{\varepsilon}}{\partial\theta_c}\delta\theta_c\,\mathrm{d}V+\int_{\Gamma_c}\boldsymbol{t}\frac{\partial[\![\boldsymbol{u}]\!]}{\partial\theta_c}\delta\theta_c\,\mathrm{d}A&=0,\\ \int_{\Omega\setminus\Gamma_c}\boldsymbol{\sigma}:\frac{\partial\boldsymbol{\varepsilon}}{\partial A_c}\delta A_c\,\mathrm{d}V+\int_{\Gamma_c}\boldsymbol{t}\frac{\partial[\![\boldsymbol{u}]\!]}{\partial A_c}\delta A_c\,\mathrm{d}A&=0.\end{aligned} \tag{7}$$

In the implementation of the model (see also Section 3.3), the derivatives with respect to the angle θ_c and the surface A_c of new crack segments are computed numerically as [1]

$$\begin{aligned}\frac{\partial\mathcal{W}_c}{\partial A_c}&=\frac{1}{\Delta A_c}\left[\mathcal{W}_c(A_c+\Delta A_c)-\mathcal{W}_c(A_c)\right],\\ \frac{\partial\mathcal{W}_c}{\partial\theta_c}&=\frac{1}{\Delta\theta_c}\left[\mathcal{W}_c(\theta_c+\Delta\theta_c)-\mathcal{W}_c(\theta_c)\right].\end{aligned} \tag{8}$$

3 Energy-Based Extended Finite Element Method

This section contains the enrichment strategy using the Extended Finite Element Method, the traction separation law employed for mixed mode conditions and a concise description of the energy-based X-FEM formulation as proposed in [1] for 2D analyses of cohesive cracks.

3.1 Enhanced Displacement Approximation

If a crack fully crosses an element, the Sign function according to Equation (3) is used to represent the discontinuous displacement field across the crack. Using standard bi-linear finite element shape functions as a partition of unity, the approximation of the displacement field within this element is - according to Equation (3) - given by a continuous part $\bar{\boldsymbol{u}}$ and a discontinuous part $\check{\boldsymbol{u}}$

$$\boldsymbol{u}=\bar{\boldsymbol{u}}+\check{\boldsymbol{u}}=\bar{\boldsymbol{u}}+S_c\hat{\boldsymbol{u}}\quad\text{with}\quad\bar{\boldsymbol{u}}\approx\sum_{i=1}^{n_r}N_i\,\bar{\boldsymbol{u}}_i^r\quad\text{and}\quad\hat{\boldsymbol{u}}\approx\sum_{i=1}^{4}N_i\,\hat{\boldsymbol{u}}_i^s, \tag{9}$$

where n_r is the number of nodes used for the spatial discretization of the regular displacement field (superscript r). In the present implementation, hierarchical higher order shape functions, allowing to chose polynomials in the

range between $p = 1$ to $p = 4$, are used for the approximation of the regular displacements. For the displacements $\hat{\boldsymbol{u}}$ enriched by the Sign-function (superscript s) four low order (bi-linear) shape functions are used.

In the vicinity of a crack tip, crack tip enhancement functions are used. As in the previous case, the approximation within elements located in the vicinity of the crack tip (see Figure 3) consists of a continuous part $\bar{\boldsymbol{u}}$ and a discontinuous part $\check{\boldsymbol{u}}$:

$$\begin{aligned} &\boldsymbol{u} = \bar{\boldsymbol{u}} + S_c \hat{\boldsymbol{u}}, \quad \text{with} \\ &\bar{\boldsymbol{u}} \approx \sum_{i=1}^{n_r} \boldsymbol{u}_i^r N_i + \sum_{i=1}^{4} \sum_{k=1}^{3} \bar{\boldsymbol{u}}_{ki}^t N_i F_k \qquad \text{and} \qquad \hat{\boldsymbol{u}} \approx \sum_{i=1}^{4} \hat{\boldsymbol{u}}_{4i}^t N_i \tilde{F}_4 \end{aligned} \tag{10}$$

where $S_c \tilde{F}_4 = F_4$ for notation consistency and again low order shape functions are used for the enriched part (superscript t). The crack tip functions for cohesive cracks F_1-F_4 are assumed as [1]

$$\begin{aligned} F_1(r_c, \theta_c) &= r_c \cos\left(\frac{\theta_c}{2}\right), \quad F_2(r_c, \theta_c) = r_c \sin\left(\frac{\theta_c}{2}\right) \sin(\theta_c), \\ F_3(r_c, \theta_c) &= r_c \cos\left(\frac{\theta_c}{2}\right) \sin(\theta_c), \quad F_4(r_c, \theta_c) = r_c \sin\left(\frac{\theta_c}{2}\right), \end{aligned} \tag{11}$$

which leads to bounded derivatives and, consequently, to bounded stresses at the crack tip (Figure 2).

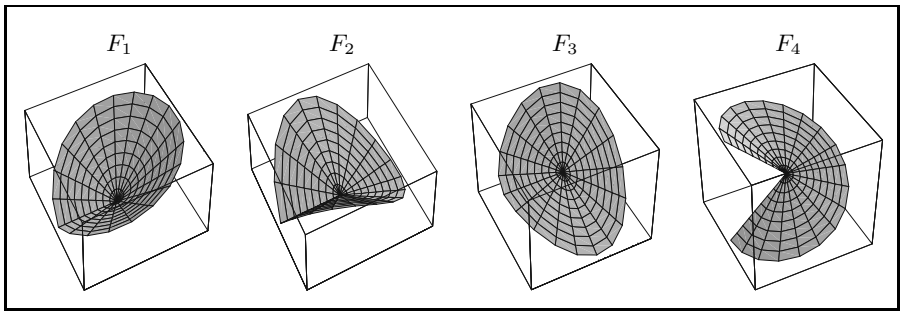


Fig. 2. Functions F_1–F_4 used for the approximation of the displacement field in the vicinity of the tip of cohesive cracks

Figure 3 illustrates the nodal enrichment strategy for open cracks and for crack tip segments according to Equations (9) and (10), respectively. Since, according to the proposed energy-based crack propagation criterion, the orientation and the length of new segments is not known a priori, a larger region in the vicinity of the root of the crack tip segment is enriched with the crack tip functions. This enhancement strategy avoids a change of the nodal enrichment

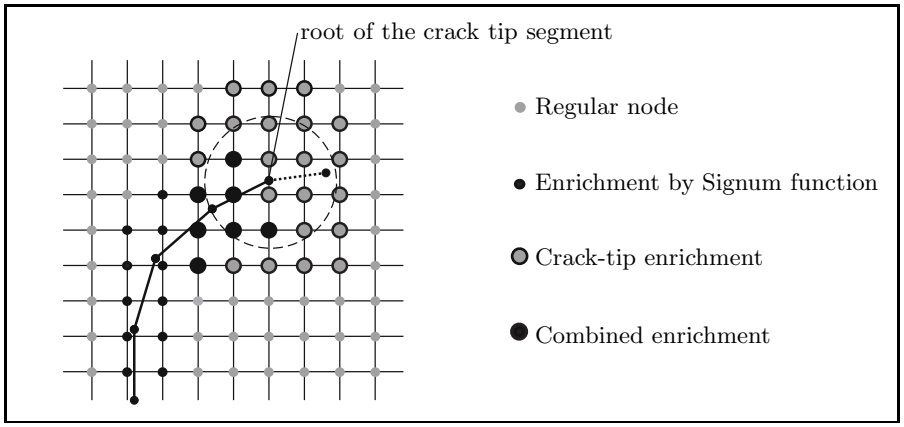


Fig. 3. Nodal enrichment strategy for open cracks and for the region in the vicinity of the crack tip

during the equilibrium iteration. The evaluation of the area integrals within elements crossed by a crack is based on a Delaunay sub-triangularization of the respective areas on both sides of the crack [12].

3.2 Traction-Separation Law for Mixed-Mode Conditions

The traction-separation law suitable for cohesive cracks in quasi-brittle materials subjected to general Mixed-Mode conditions follows the formulation introduced by [13] and later employed in [5]. It accounts for dissipative interface mechanisms in Mode-I and Mode-II conditions using an equivalent traction separation law.

The relation between the traction vector $\boldsymbol{t}$ and the displacement jump $[\![\boldsymbol{u}]\!]$ representing its energetic conjugate variable is given in the general format

$$\boldsymbol{t} = \mathbb{T}\,[\![\boldsymbol{u}]\!] = \left(\mathbb{T}^e - \mathbb{T}^d\right)[\![\boldsymbol{u}]\!]\,, \tag{12}$$

where $\mathbb{T}^e = T^e \boldsymbol{I}$ is an initial isotropic (elastic) stiffness corresponding to the initial uncracked situation and $\mathbb{T}^d$ denotes a damage tensor governing the degradation of the stiffness $\mathbb{T}$.

In the present model, the relation (12) is formulated in terms of an equivalent traction t_{eq} and an equivalent crack opening $[\![u]\!]_{eq}$, respectively

$$t_{eq} = \left(T^e - T^d\right)[\![u]\!]_{eq}\,, \tag{13}$$

with $[\![u]\!]_{eq}$ defined as:

$$[\![u]\!]_{eq} = \sqrt{[\![u]\!]_n^2 + \beta^2\,[\![u]\!]_s^2}. \tag{14}$$

The subscripts n and s refer to normal and shear components, respectively. The parameter β governs the ratio between the strength and stiffness in mode I and mode II opening. For both modes of fracture the same fracture energy G_c is assumed in the model (see [1, 13] for more details).

3.3 Finite Element Formulation

According to the proposed energy-based crack propagation model [1] the angle θ_c and the length r_c of new crack segments are introduced as additional degrees of freedom in the discretized structural model. Accordingly, the stationarity condition (6) and its linearization is recast in vector- and matrix-form, respectively, in the format

$$\begin{aligned}
\delta\Psi &= \begin{bmatrix} \delta\boldsymbol{u} \\ \delta r_c \\ \delta\theta_c \end{bmatrix} \cdot \left[\begin{bmatrix} \boldsymbol{r}^u \\ r^r \\ r^\theta \end{bmatrix} - \begin{bmatrix} \boldsymbol{r}^{\star u} \\ 0 \\ 0 \end{bmatrix} \right] = 0, \\
\Delta\delta\Psi &= \begin{bmatrix} \delta\boldsymbol{u} \\ \delta r_c \\ \delta\theta_c \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{k}^{uu} & \boldsymbol{k}^{ur} & \boldsymbol{k}^{u\theta} \\ \boldsymbol{k}^{ru} & k^{rr} & k^{r\theta} \\ \boldsymbol{k}^{\theta u} & k^{\theta r} & k^{\theta\theta} \end{bmatrix} \begin{bmatrix} \Delta\boldsymbol{u} \\ \Delta r_c \\ \Delta\theta_c \end{bmatrix},
\end{aligned} \tag{15}$$

where $\boldsymbol{r}^{\star u}$ is the vector of external forces. The vectors of internal forces are identified from (6) as

$$\begin{aligned}
\boldsymbol{r}^u &= \partial\mathcal{U}/\partial\boldsymbol{u} + \partial\mathcal{W}_c/\partial\boldsymbol{u}, \\
r^\theta &= \partial\mathcal{U}/\partial\theta_c + \partial\mathcal{W}_c/\partial\theta_c, \\
r^r &= \partial\mathcal{U}/\partial r_c + \partial\mathcal{W}_c/\partial r_c.
\end{aligned} \tag{16}$$

The computation of the components of the vectors $\boldsymbol{r}^r$, $\boldsymbol{r}^\theta$, and of the submatrices $\boldsymbol{k}^{ur}$, $\boldsymbol{k}^{u\theta}$, $\boldsymbol{k}^{rr}$, $\boldsymbol{k}^{r\theta}$ and $\boldsymbol{k}^{\theta\theta}$ requires the determination of the first and the second derivatives of the total energy (1) with respect to the crack angle θ_c and the segment length r_c. Due to the complexity involved in the analytical determination of these derivatives, they are computed by means of numerical differentiation. Details of the numerical determination of these components are contained in [1, 12].

The incremental-iterative solution of boundary value problems using the proposed energy-based Extended finite element method is characterized by applying the total load in a sequence of incremental steps $[t_n, t_{n+1}]$ and enforcing the the stationarity condition (15_1)

$$\delta\Psi_{n+1}(\mathbf{u}, r_c, \theta_c) = 0 \tag{17}$$

by solving (17) simultaneously for the three unknowns $\mathbf{u}$, r_c and θ_c. The format of (17) and its algorithmic formulation is identical to multifield problems. The starting value of the crack angle $\theta_{c,n+1}$ is always taken according to the previous converged state $\theta_{c,n}$.

If a crack is signalled to propagate, the coupled equation system

$$\begin{bmatrix} \boldsymbol{r}^u \\ r^r \\ r^\theta \end{bmatrix}^k_{n+1} - \begin{bmatrix} \boldsymbol{r}^{\star u} \\ 0 \\ 0 \end{bmatrix}_{n+1} + \begin{bmatrix} \delta\boldsymbol{u} \\ \delta r_c \\ \delta\theta_c \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{k}^{uu} & \boldsymbol{k}^{ur} & \boldsymbol{k}^{u\theta} \\ \boldsymbol{k}^{ru} & k^{rr} & k^{r\theta} \\ \boldsymbol{k}^{\theta u} & k^{\theta r} & k^{\theta\theta} \end{bmatrix} \begin{bmatrix} \Delta\boldsymbol{u} \\ \Delta r_c \\ \Delta\theta_c \end{bmatrix}^{k+1}_{n+1} = 0, \quad (18)$$

is solved within the k-th iteration of a loading step $[t_n, t_{n+1}]$. The stationary value of the energy functional provides, for the given increment of external load, the correct crack configuration. After convergence, the state of displacements $\mathbf{u}_{n+1}$, crack angle $\theta_{c,n+1}$ and crack length $r_{c,n+1}$ is updated and the external load is increased.

4 Influence of Integration

The robustness of the proposed energy-based X-FEM model using the NEWTON's method to solve the coupled system of equations (18) highly depends on the smoothness of the solution. This subsection contains a numerical study on the influence of the mode of integration on the energy distribution.

To this end, the classical example of a crack growing from a fillet of a plane T-shaped structure is considered. The geometry, loading and boundary conditions of the investigated structure together with its discretization are illustrated in Figure 4a. Linear elastic material behavior with Young's modulus E=200 [GPa], Poisson's ratio ν=0.3 and plane strain conditions are assumed. The thickness of the slab is d=1 [m] and the applied tension is σ=1 [KPa].

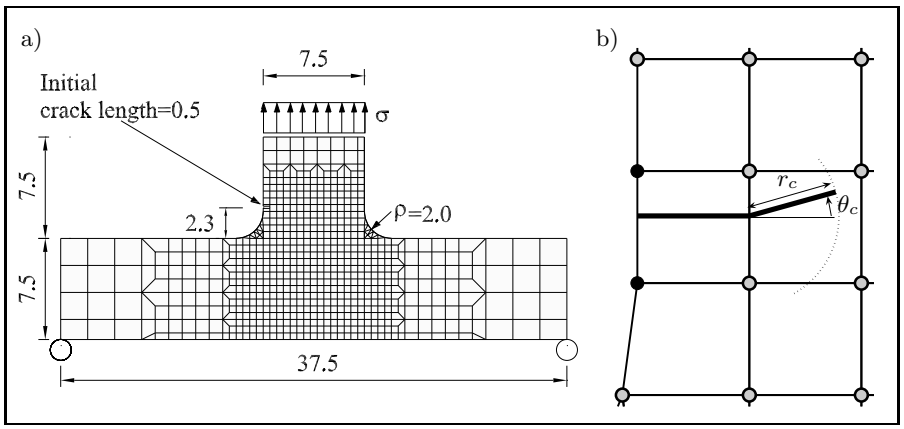


Fig. 4. Numerical study on the influence of integration on the energy distribution: a) Geometry, loading and discretization of the benchmark example, b) Crack tip zone (Units in [cm])

In the numerical study, only one loading step is considered. Restricting the attention to a region ahead of the initial crack tip as depicted in Figure 4b, the distribution of the total energy as a function of the crack kinking angle θ_c of a new crack segment r_c with fixed length (r_c=0.4 [cm]), covering a range -60$\leq \theta_c \leq$ 40, is investigated. The crack tip element is a square with size 0.5 [cm] (Figure 4). The nodes with near tip and combined enrichment are marked according to Figure 3. Since, according to Subsection 3.1, all nodes located within a circular area in the vicinity of the crack tip are enriched, unlike the case where just the element containing the crack tip would be enriched, no jumps in the global energy distribution are expected if an exact integration is carried out.

The following situations are identified as potential sources for discontinuities in the energy distribution:

Cases A and C: The crack tip changes from one element to another. This is associated with an abrupt change of the order of integration of the adjacent element. In the investigated benchmark, this occurs for $\theta_c \approx$ -49(Case A) and $\theta_c \approx$ 30(Case C) see Figure 5.

The Delaunay triangulation used for the subintegration of the crack tip element changes abruptly. This occurs at $\theta_c \approx 16$, see Figure 6.

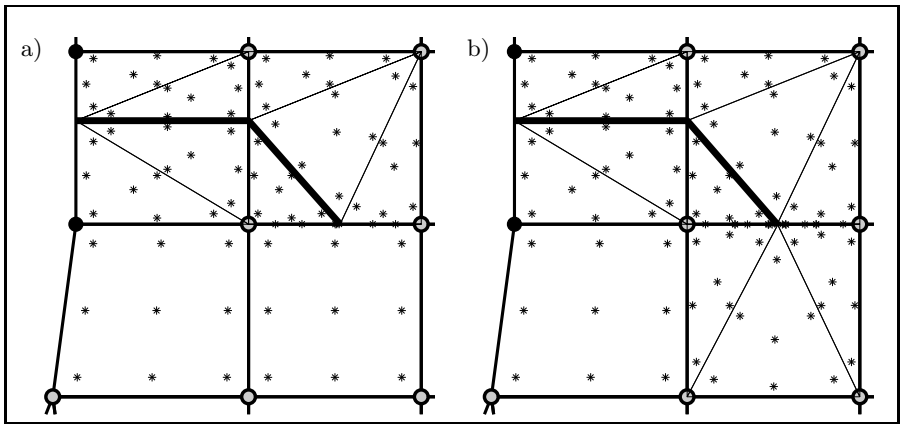


Fig. 5. Numerical study on the influence of integration on the energy distribution: a) Crack tip located within original crack tip element, b) Crack tip changes to neighboring element

Figures 7 and 8 show the global energy $\Psi(\mathbf{u}, \theta_c)$ as a function of the crack kinking angle θ_c for different orders of integration. The total energy Ψ has been normalized with respect to the absolute value of the energy corresponding to the configuration without crack extension Ψ_0. The three abovementioned cases A, B and C are marked in these figures.

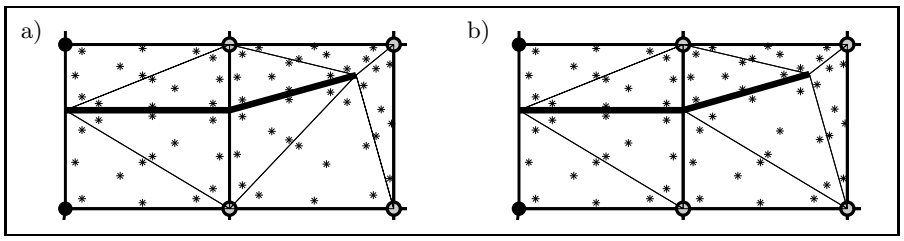


Fig. 6. Numerical study on the influence of integration on the energy distribution: Abrupt change of Delaunay triangulation within the crack tip element (Case B)

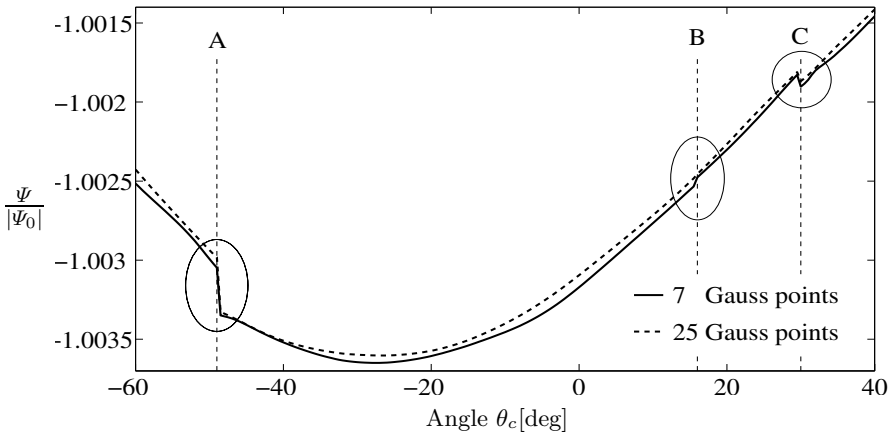


Fig. 7. Normalized global energy vs. crack angle θ_c considering 7 and 25 Gauss points for the subtriangles and 2× 2 integration for the quadrilaterals

Figure 7 refers to analyses using a 2×2 integration of the finite elements. As far as the integration of the subtriangulation in the crack tip element is concerned, an integration scheme using 7 integration points is compared with a scheme using 25 integration points. The three states of discontinuities are also marked in Figure 7. This figure shows, that the discontinuity related to Case B can be easily resolved using a higher order of integration (25 instead of 7 Gauss points) for the subtriangles. However, the discontinuities associated with A and C are not resolved.

To this end, an alternative scenario, characterized by a higher order integration of the finite elements using a 40×40 instead of 2×2 integration scheme has been investigated. Figure 7 shows, that with a higher order integration of the neighboring (uncracked) elements, the discontinuity connected with the crossing of the crack tip from one element to the other (stages A and C) is substantially reduced. It is concluded, that a comparable level of integration order is required in all elements located in the vicinity of the crack tip.

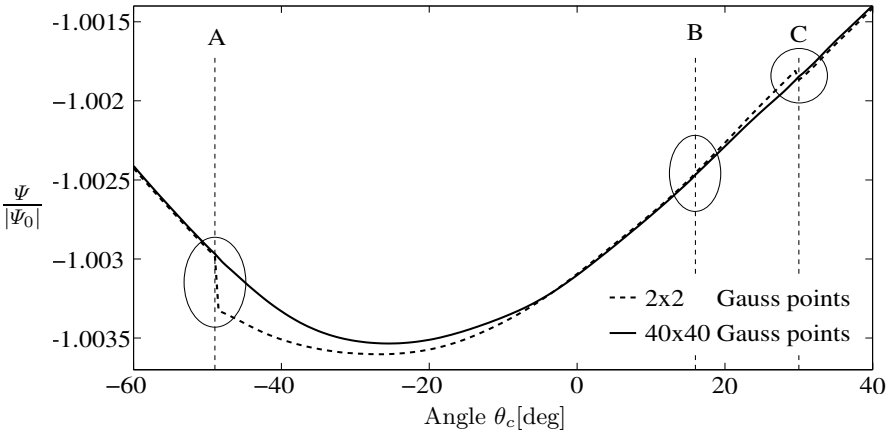


Fig. 8. Normalized global energy vs. crack angle θ_c considering 2×2 and 40×40 integration for the quadrilaterals and 25 Gauss points for the subtriangles

5 Influence of the Interface Law

The influence of the interface law on the predicted crack trajectory is investigated by means of re-analyses of two classical benchmark examples for cohesive crack propagation: A L-shaped panel made of concrete, characterized primarily by Mode-I fracture (Subsection 5.1) and a notched panel made of mortar subjected to mixed mode condition (Subsection 5.2). In addition, the influence of taking different fracture energies in mode I and mode II conditions on the crack propagation direction within a biaxially compressed slab is investigated in Subsection 5.2.

5.1 Crack Propagation in a L-Shaped Panel

The first numerical example is a L-shaped panel, which has been tested experimentally by [14] and has been used as a benchmark problem by several researchers [14–16]. The geometry, the loading and boundary conditions of the panel as well as the material parameters are contained in Figure 9.

The major interest of this study is to investigate the influence of the shear transfer law, governed by the coupling parameter β, on the predicted crack path. Three values of the parameter β, representing different values of the elastic shear stiffness as well as of the shear strength, have been used for the analyses: $\beta = 1$ (small shear stiffness and strength), $\beta = 1.5$ (medium shear stiffness and strength) and $\beta = 2$ (large shear stiffness and strength). It should be noted, that no attempt was made to calibrate the value of β directly from experiments on cracked specimens.

Figure 10 contains the numerical results obtained by means of the damaging interface model for three different values of the shear parameter β. This parameter represents, on a macroscopic level, the load carrying capacity of the

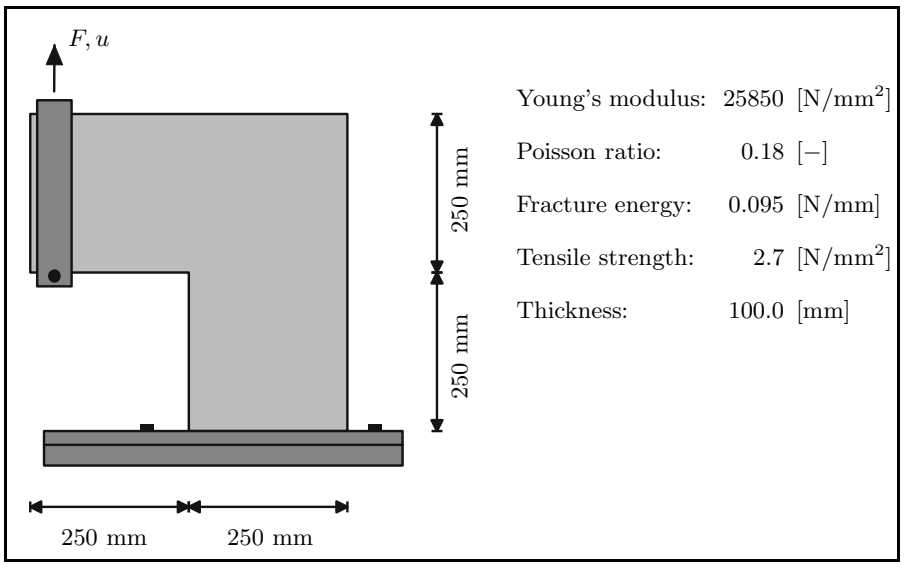


Fig. 9. Numerical analysis of a L-shaped panel: Geometry and material parameters

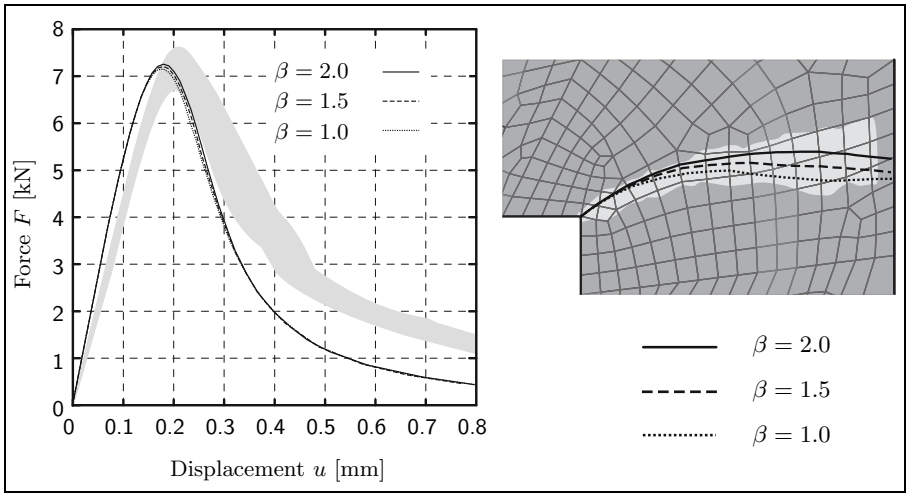


Fig. 10. Numerical analysis of a L-shaped panel: Computed crack paths obtained from the damage interface law using three different values for the parameter β

crack faces, which depends on the asperities and the tortuosity of the crack, and, consequently, on the microstructure of the material. As is illustrated in this Figure, the shear resistance expressed by the parameter β has a noticeable influence on the predicted crack path. The crack path obtained from a value of $\beta = 1.5$ lies reasonably well within the range of the experimental

results. Also the crack trajectory seems to agree well with the median of the laboratory observations. Despite these noticeable differences, all three crack paths are located within the range of experimental results indicated in grey color in Figure 10.

5.2 Crack Propagation in a Tension-Shear Specimen

A benchmark problem characterized by Mixed-Mode crack propagation in a concrete specimen which has been investigated experimentally by Nooru--Mohammed [17] is investigated numerically. In this test, a notched panel (see Figure 11) made of mortar has been first subjected to a shear force F_s= 10 [kN] applied along the upper left and the lower right side of the specimen. Subsequentially, displacement controlled tensile loading F_n was applied on the upper face of the specimen. For additional analyses of this benchmark test by means of the energy-based X-FEM model we refer to [1].

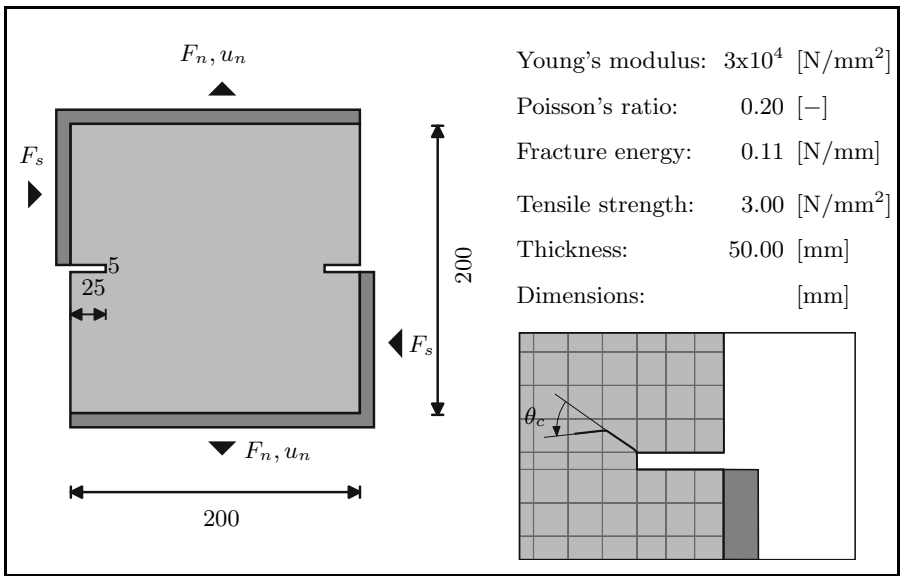


Fig. 11. Numerical analysis of a tension-shear test: Geometry and material parameters

The finite element discretization consists of 435 elements. Quadratic shape functions are used for the approximation of the regular displacements $\bar{\boldsymbol{u}}$ (see Equation (9)) while bi-linear functions are used as Partition of Unity.

The material parameters used for this study are also contained in Figure 11. It should be noted that, as in the previous example, no attempt has been made to adjust the parameters to the numerical results. Since no inform-

ation on the fracture energy G_f is contained in [17], the value of G_f has been taken from [18].

First, the influence of the shear transfer model, i.e. the influence of the value of the coupling parameter β on the distribution of the total energy Ψ with a variation of the crack propagation angle θ_c is investigated. To this end, only one single loading step is considered. Starting from an initial crack segment, a new crack segment with constant length r_c associated with an increment of the prescribed load is assumed to propagate from the existing crack tip at different angles θ_c within a range of $\theta_c = -22.9°$ to $\theta_c = 31.5°$ (see Figure 11). For each position of the new segment an equilibrium state was computed and the respective values of the total energy Ψ and of the surface energy $\mathcal{W}_c$ were recorded. Figure 12a contains the respective diagrams of the total energy Ψ versus the crack kinking angle θ_c within the investigated interval. As is noted from this Figure, the energy functional is not convex for all investigated situations. For $\beta = 2.0$ and $\beta = 1.5$, the functional Ψ is locally convex with a minimum at $\approx 1.9^o$ and $\approx 5.8^o$, respectively. This range of the crack kinking angles corresponds well with experimental observations.

For $\beta = 1$, i.e. assuming a small shear transfer capacity along the interface, no stationarity point is existing within the investigated range of crack kinking angles. Consequently, using the proposed energy-based crack model, which is based on the simultaneous solution of $\delta\Psi(\boldsymbol{u}, \theta_c, r_c)$ no converged solution could be found for $\beta = 1$. Also, for $\beta = 1.5$ convergence problems have been observed and the analysis could not be completed.

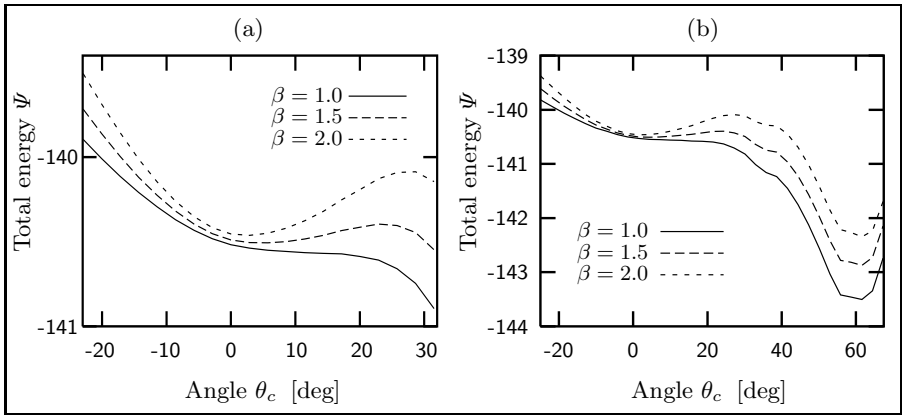


Fig. 12. Numerical analysis of tension-shear benchmark problem: Distribution of total energy Ψ [Nm] versus the crack propagation angle θ_c within a range of a) -23^o $\leq \theta_c \leq$ 31.5^o and b) -23^o $\leq \theta_c \leq$ 70^o

Figure 12b, shows the energy distribution for a larger range of θ_c (-23^o $\leq \theta_c \leq$ 70^o) computed for the three values of β. According to this figure, a minimum energy level exists at $\theta_c \approx 60^o$ for all three analyses. This second

energy minimum corresponds to a Mode II-dominated crack propagation. It is assumed, that this second energy minimum would not appear if a more realistic interface model, characterized by different values of the fracture energy in Mode I and Mode II crack opening, is used. This assumption is corroborated by the results presented in the following subsection.

5.3 Crack Propagation in a Slab Subjected to Biaxial Compression

In the interface model used in the energy-based X-FEM model the same value of the fracture energy in mode-I and mode-II is assumed, which, as alluded above, strongly influences the energy distribution and the location of local and global minima of Ψ.

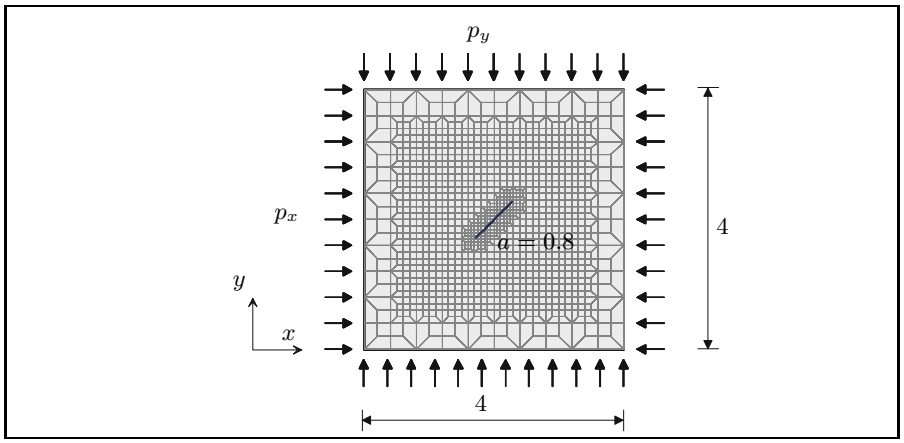


Fig. 13. Crack propagation from an existing inclined crack in a slab subjected to compression: Geometry and discretization (Units [cm])

In this subsection, numerical analyses of the propagation of an inclined flaw located in a plane slab subjected to biaxial compression (Figure 13) are performed in order to study the distribution of the total energy Ψ as a function of the kinking angle θ_c and to assess the relevance of taking different values of the fracture energy in mode-I and mode-II conditions. The existing crack has the length of 0.8 cm and is inclined 45^o to the loading directions. The Young's modulus and the Poisson's ratio are taken as E= 1000 MPa, ν=0.28 In these analyses, no cohesion was assumed in the crack faces, i.e. linear elastic fracture mechanics analyses have been performed.

Figure 14 illustrates the distribution of the total energy Ψ as a function of the kinking angle θ_c of a new crack segment with r_c= 0.08 cm for different ratios p_y/p_x of the axial and the lateral loading. A minimum of Ψ is obtained approximatly within a range of $\theta_c \approx 20^o - 40^o$, which corresponds to

a shear-dominated crack propagation. However, in laboratory tests on sandstone, tensile wing cracks at $\theta_c \approx 90^o$ have been observed [19].

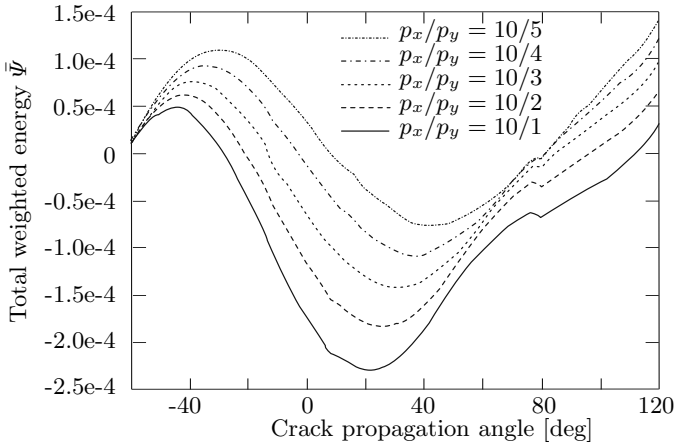


Fig. 14. Total energy Ψ as a function of the crack propagation angle θ_c for $G_{Ic}/G_{IIc} = 1$

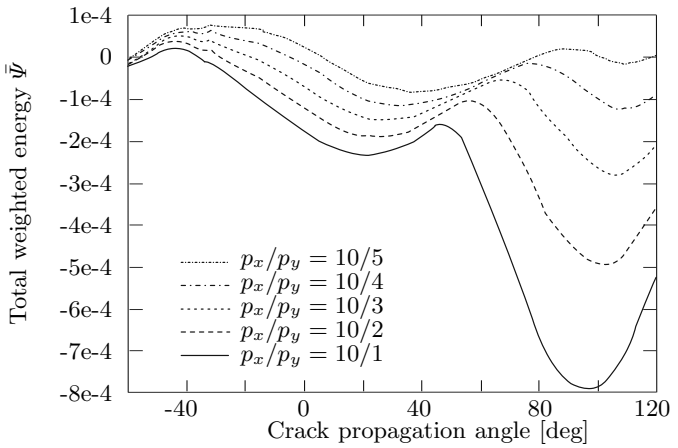


Fig. 15. Total energy Ψ as a function of the crack propagation angle θ_c for $G_{Ic}/G_{IIc} = 1/10$

Therefore, following the work of Shen & Stephansson [19], the partial energies Ψ_I and Ψ_{II} attributed to a pure mode-I (pure separation) and a pure mode-II (pure sliding) mechanism, respectively, have been computed separately and associated with different weigths to plot a weighted total energy

$\bar{\Psi}$ as a function of θ_c according to

$$\bar{\Psi} = \Psi_I + \Psi_{II} G_{Ic}/G_{IIc} \;\rightarrow\; min. \tag{19}$$

To this end, different values of the fracture energies G_{Ic} and G_{IIc} associated with mode I and mode II conditions, respectively, are assumed. For $G_{Ic} = G_{IIc}$ the total energy $\bar{\Psi}$ is given by the sum $\Psi_I + \Psi_{II}$. This case is illustrated in Figure 14.

Figure 15 contains the plots of the total energy according to (19) using weighted fracture energies with a ratio $G_{Ic}/G_{IIc} = 1/10$. Indeed, now a local minimum of the energy is observed at the location where previously a global minimum was observed, and another global minimum now existis at $\theta_c \approx 90^o - 100^o$. This minimum corresponds, as expected, to a tensile wing crack propagating more or less orthogonal to the direction of the existing crack.

6 Concluding Remarks

This paper contains numerical studies of a recently proposed energy-based approach to model propagation of cohesive cracks in quasi-brittle materials within the framework of the Extended Finite Element Method [1]. The main focus has been laid on selected aspects of the model expecting to have an impact on the energy distribution and, consequently, on the robustness and the reliability of the results. According to this model cracks are allowed to continuously propagate through elements and the crack propagation angle θ_c and the surface of new crack segments A_c are determined from a variational formulation characterized by the stationarity of the total energy $\Psi(\boldsymbol{u}, \theta_c, A_c)$ of the structure. Any ad hoc assumptions often employed for the determination of the crack propagation angle are avoided. The finite element formulation has the typical structure of a multifield problem, with only two additional unknowns to be introduced for each propagating crack.

It was shown, that the order of integration within the cracked as well as within the neighboring elements in the vicinity of the crack tip has a considerable influence on the smoothness of the energy distribution, and, consequently, on the robustness of the numerical solution using NEWTON'S method. It was also shown, that the shear transfer capacity represented by the interface model has a considerable influence on the distribution of the energy for varying crack kinking angles θ_c and, consequently, on the location of the minimum of this (locally convex) distribution. As was shown by numerical analyses of mixed-mode crack propagation in a tension-shear specimen, physically unrealistic assumptions for the interface law are reflected by the fact that local convexity of the energy distribution $\Psi(\theta_c)$ may be lost within the expected range of (physically plausible) crack kinking angles. According to the interface law, in which the same fracture energies are used in mode I and mode II conditions, alternative modes of crack propagation are indicated

by energy minima at larger kinking angles corresponding to mode II fracture. The necessitiy of taking different fracture energies in mode I and mode II conditions into account has been corroborated by numerical analyses of the propagation of a wing crack in a slab subjected to biaxial compression.

Acknowledgement

The authors thank Prof. G. Dresen and Dr. A. Zang from the Potsdam Research Center for Geosciences for partially supporting the second author.

References

1. G. Meschke and P. Dumstorff. Energy-based modeling of cohesive and cohesionless cracks via X-FEM. *Computer Methods in Applied Mechanics and Engineering*, 2007. in press.
2. N. Moës, J.E. Dolbow, and T. Belytschko. A finite element method for crack growth without remeshing. *International Journal for Numerical Methods in Engineering*, 46:131–150, 1999.
3. N. Moës and T. Belytschko. Extended finite element method for cohesive crack growth. *Engineering Fracture Mechanics*, 69:813–833, 2002.
4. G.N. Wells and L.J. Sluys. A new method for modelling cohesive cracks using finite elements. *International Journal for Numerical Methods in Engineering*, 50:2667–2682, 2001.
5. S. Mariani and U. Perego. Extended finite element method for quasi-brittle fracture. *International Journal for Numerical Methods in Engineering*, 58:103–126, 2003.
6. J. Mergheim, E. Kuhl, and P. Steinmann. A finite element method for the computational modelling of cohesive cracks. *International Journal for Numerical Methods in Engineering*, 63:276–289, 2005.
7. J. Oliver, A.E. Huespe, E. Samaniego, and E.W.V. Chaves. On strategies for tracking strong discontinuities in computational failure mechanics. In H.A. Mang, F.G. Rammerstorfer, and J. Eberhardsteiner, editors, *Fifth World Congress on Computational Mechanics (WCCM V)*. Online publication, 2002.
8. J. Oliver, A.E. Huespe, E. Samaniego, and E.W.V. Chaves. Continuum approach to the numerical simulation of material failure in concrete. *International Journal for Numerical and Analytical Methods in Geomechanics*, 28:609–632, 2004.
9. G.A. Francfort and J.J. Marigo. Revisiting brittle fracture as an energy minimization problem. *Journal of Mechanics and Physics of Solids*, 46(8):1319–1342, 1998.
10. P. Dumstorff and G. Meschke. Crack propagation criteria in the framework of X-FEM-based structural analyses. *International Journal for Numerical and Analytical Methods in Geomechanics*, 31, 2007. in press.
11. K. C. Le. Variational principles of non-linear theory of brittle fracture mechanics. *Journal of Applied Mathematics and Mechanics*, 54:543–549, 1990.

12. P. Dumstorff. *Modellierung und numerische Simulation von Rissfortschritt in spröden und quasi-spröden Materialien auf Basis der Extended Finite Element Method*. PhD thesis, Institute for Structural Mechanics, Ruhr University Bochum, 2005. in german.
13. G.T. Camacho and M. Ortiz. Computational modelling of impact damage in brittle materials. *International Journal of Solids and Structures*, 33:2899–2938, 1996.
14. B. J. Winkler. *Traglastuntersuchungen von unbewehrten und bewehrten Betonstrukturen auf der Grundlage eines objektiven Werkstoffgesetzes für Beton*. PhD thesis, Universität Innsbruck, 2001.
15. J. Mosler and G. Meschke. Embedded cracks vs. smeared crack models: A comparison of elementwise discontinuous crack path approaches with emphasis on mesh bias. *Computer Methods in Applied Mechanics and Engineering*, 193:3351–3375, 2004. (Special issue on Computational Failure Mechanics).
16. C. Feist and G. Hofstetter. Mesh-insensitive strong discontinuity approach for fracture simulations of concrete. In *Numerical Methods in Continuum Mechanics (NMCM 2003)*, 2003.
17. M.B. Nooru-Mohamed. *Mixed-mode Fracture of Concrete: an Experimental Approach*. PhD thesis, Technische Universiteit Delft, 1992.
18. C. Feist. *A Numerical Model for Cracking of Plain Concrete Based on the Strong Discontinuity Approach*. PhD thesis, University Innsbruck, 2004.
19. B. Shen and O. Stephansson. Modification of the g-criterion for crack propagation subjected to compression. *Engineering Fracture Mechanics*, 47:177–189, 1994.

An Evaluation of the Accuracy of Discontinuous Finite Elements in Explicit Dynamic Calculations

Joris J.C. Remmers[1], René de Borst[1,2], and Alan Needleman[3]

[1] Delft University of Technology, Faculty of Aerospace Engineering, PO Box 5058, 2600 GB Delft, The Netherlands; J.J.C.Remmers@TUDelft.nl
[2] LaMCoS – UMR CNRS 5514, INSA de Lyon, F-69021 Villeurbanne, France; R.deBorst@TUDelft.nl
[3] Division of Engineering, Brown University, Providence, Rhode Island 02912, USA; Alan_Needleman@brown.edu

Summary. The use of partition of unity based discontinuous finite element formulations for the simulation of crack propagation with implicit simulations is now well established. However, in explicit simulations the accuracy is still a point of concern. Some outstanding issues will be addressed in this paper.

Key words: Explicit time integration, fast crack growth, partition of unity method

1 Introduction

Over the last years, partition of unity based finite element formulations have become one of the most important approaches to simulate fracture processes. Initially, the technique served as an elegant way to represent the traction-free crack and the singular displacement field in linear elastic fracture mechanics models [1]. Later, the approach has been used together with a cohesive zone formulation for the simulation of fracture in brittle materials [2, 3] and, in combination with geometrically nonlinear kinematic relations, to analyse delamination growth in composite materials [4, 5].

Recent developments have focused on the application of the approach to the simulation of fast crack growth in brittle solids [6–8]. Since the speed of the propagating crack tip is in the order of magnitude of the Rayleigh wave speed, explicit solution algorithms are usually preferred for these analyses. In contrast to implicit methods, the convergence of the solution is not guaranteed, which may lead to inaccurate results.

In this paper, the performance of discontinuous elements in explicit simulations is investigated for three different cases. First, the effect of lumping

Alain Combescure et al. (eds.), IUTAM Symposium on Discretization Methods for Evolving Discontinuities, 303–322.

the mass matrix is studied. Lumping the mass matrix is a standard procedure to decouple the system of equations and to speed up the simulations. In this case, however, important information regarding the coupling of the regular and enhanced degrees of freedom can be lost, which may influence the results. In a second benchmark test, the accuracy of the Newmark-β method is tested. As a rule of thumb, the critical time-step in conventional finite element simulations is equal to the dilatational speed of the material divided by the specific element length. When using discontinuous elements, it is questionable whether this specific length is equal to the original length of the element. The specific length can then be equal to the size of the smallest part of an element that is crossed by a discontinuity, which will have large implications on the critical time step. Finally, the effects of a propagating discontinuity are investigated. Adding degrees of freedom to a system may affect the energy balance of the system. Because the convergence of the calculation is not guaranteed this can lead to spurious energy dissipation, or even worse, the creation of energy. In this test, a cohesive constitutive relation for the crack opening is used.

2 Finite Element Formulation

Kinematically, cracks are represented as a discontinuity in the displacement field. The body Ω in Figure 1 is crossed by a discontinuity Γ_{d}. The two parts on either side of Γ_{d} are denoted by Ω^+ and Ω^-, respectively. The displacement of a material point on this body with coordinate $\mathbf{x}$ at time t can be written as a regular displacement field $\hat{\mathbf{u}}$ plus the additional field $\tilde{\mathbf{u}}$ according to [1, 2]:

$$\mathbf{u}(\mathbf{x},t) = \hat{\mathbf{u}}(\mathbf{x},t) + \mathcal{H}_{\Gamma_{\mathrm{d}}}(\mathbf{x})\tilde{\mathbf{u}}(\mathbf{x},t)\,, \tag{1}$$

where $\mathcal{H}_{\Gamma_{\mathrm{d}}}$ represents a step function which has constant but different values on either side of the discontinuity:

$$\mathcal{H}_{\Gamma_{\mathrm{d}}} = \begin{cases} \mathcal{H}^+ & \text{if} \quad \mathbf{x} \in \Omega^+, \\ \mathcal{H}^- & \text{if} \quad \mathbf{x} \in \Omega^-\,. \end{cases} \tag{2}$$

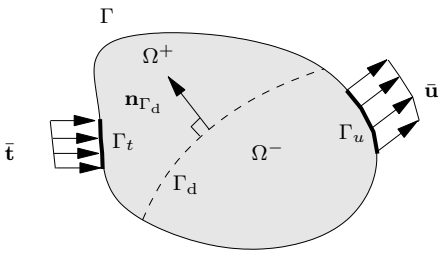


Fig. 1. A domain Ω is crossed by a discontinuity Γ_{d} (denoted by the dashed line).

Table 1. An overview of different step functions that can be used to create a jump in the displacement field, see equation (2).

Jump function	$\mathcal{H}^+$	$\mathcal{H}^-$	h
Heaviside	1	0	1
Symmetric	1/2	-1/2	1
Unit Symmetric	1	-1	2

In most cases, a standard Heaviside function is used [2], but alternative functions can be considered as well, e.g. [7]. An overview of possible step functions is given in Table 1.

The opening of the crack is equal to the magnitude of the jump in the displacement field at the discontinuity and is given by:

$$\mathbf{v} = h\tilde{\mathbf{u}} \tag{3}$$

with the constant h defined in Table 1. The velocity and acceleration of a material point can be found by differentiating the displacement field, equation (1) with respect to time:

$$\dot{\mathbf{u}} = \dot{\hat{\mathbf{u}}} + \mathcal{H}_{\Gamma_\mathrm{d}}\dot{\tilde{\mathbf{u}}}\,; \qquad \ddot{\mathbf{u}} = \ddot{\hat{\mathbf{u}}} + \mathcal{H}_{\Gamma_\mathrm{d}}\ddot{\tilde{\mathbf{u}}}\,; \tag{4}$$

where $(\dot{})$ and $(\ddot{})$ denote the first and second time derivative, respectively.

The displacement and velocity fields can be used to determine the strain (rates) in a material point, which in combination with a constitutive relation gives the local the Cauchy stress $\boldsymbol{\sigma}$. The presence of the discontinuity in the displacement field does not impose any restrictions on the choice of the strain formulation. Moreover, there is no restriction on the choice of constitutive relation.

The acceleration and the stress field can be put in the linear momentum balance, which reads:

$$\rho\ddot{\mathbf{u}} + \nabla \cdot \boldsymbol{\sigma} = \mathbf{0}\,, \tag{5}$$

where ρ denotes the density of the material. The balance equation can be expressed in a weak form by multiplication with an admissible variational displacement field $\delta\mathbf{u}$ that belongs to the same space as the actual displacement field (1), thus consisting of a regular part $\delta\hat{\mathbf{u}}$ and an additional part $\delta\tilde{\mathbf{u}}$:

$$\delta\mathbf{u} = \delta\hat{\mathbf{u}} + \mathcal{H}_{\Gamma_\mathrm{d}}\delta\tilde{\mathbf{u}}\,. \tag{6}$$

The displacement and strain fields can be cast in a discrete form by employing the partition of unity property of finite element shape functions. It has been demonstrated by Babuška and Melenk [9] that any field f, even if it is not continuous, can be discretised using the finite element shape functions in combination with a set of enhanced basis functions, according to:

$$f(\mathbf{x},t) = \sum_{i=1}^{n_{nod}} \phi_i(\mathbf{x}) \left(a_i(t) + \sum_{j=1}^{m} \gamma_j(\mathbf{x}) b_{ij}(t) \right). \tag{7}$$

In this equation ϕ_i is the shape function associated to a node i, a_i is the corresponding regular degree of freedom of that node, $\gamma_j(\mathbf{x},t)$ is an enhanced basis with m orthogonal terms and b_{ij} are the additional degrees of freedom in node i that support the enhanced basis functions j. The number m of enhanced base functions may be different for each node i in the mesh. However, in order to avoid linear dependency, the enhanced basis γ_j and the shape functions ϕ_i may not originate from the same span of functions.

In the present situation, the displacement and acceleration fields can be written as:

$$\mathbf{u} = \mathbf{N}\mathbf{a} + \mathcal{H}_{\Gamma_\mathrm{d}}\mathbf{N}\mathbf{b}\,; \qquad \ddot{\mathbf{u}} = \mathbf{N}\ddot{\mathbf{a}} + \mathcal{H}_{\Gamma_\mathrm{d}}\mathbf{N}\ddot{\mathbf{b}}\,. \tag{8}$$

where $\mathbf{N}$ is a matrix containing the finite element shape functions and $\mathbf{a}$ and $\mathbf{b}$ are the regular and additional nodal displacement components, respectively. Inserting these expressions and the corresponding discrete admissible variations, into the weak form linear momentum balance gives the following set of equilibrium equations:

$$\begin{bmatrix} \mathbf{M}_{aa} & \mathbf{M}_{ab} \\ \mathbf{M}_{ab} & \mathbf{M}_{bb} \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{a}} \\ \ddot{\mathbf{b}} \end{bmatrix} = \begin{bmatrix} \mathbf{f}_a^{\mathrm{ext}} \\ \mathbf{f}_b^{\mathrm{ext}} \end{bmatrix} - \begin{bmatrix} \mathbf{f}_a^{\mathrm{int}} \\ \mathbf{f}_b^{\mathrm{int}} \end{bmatrix}, \tag{9}$$

where $\mathbf{f}_a^{\mathrm{int}}$ and $\mathbf{f}_b^{\mathrm{int}}$ denote the internal force vectors associated with the regular and enhanced degrees of freedom, respectively and $\mathbf{f}_a^{\mathrm{ext}}$ and $\mathbf{f}_b^{\mathrm{ext}}$ the corresponding external forces. For a detailed description, see [8] The terms in the mass matrix $\mathbf{M}$ are equal to:

$$\begin{aligned} \mathbf{M}_{aa} &= \int_\Omega \rho \mathbf{N}^{\mathrm{T}}\mathbf{N}\mathrm{d}\Omega\,; \\ \mathbf{M}_{ab} &= \int_\Omega \mathcal{H}_{\Gamma_\mathrm{d}} \rho \mathbf{N}^{\mathrm{T}}\mathbf{N}\mathrm{d}\Omega\,; \\ \mathbf{M}_{bb} &= \int_\Omega \mathcal{H}_{\Gamma_\mathrm{d}}\mathcal{H}_{\Gamma_\mathrm{d}} \rho \mathbf{N}^{\mathrm{T}}\mathbf{N}\mathrm{d}\Omega\,. \end{aligned} \tag{10}$$

Note that this element mass matrix is completely filled and contains terms that couple the regular degrees of freedom $\mathbf{a}$ to the additional degrees of freedom $\mathbf{b}$.

3 Solution Procedure

The equilibrium equations are discretised in the time domain using a variant of the Newmark-β explicit time integration scheme with $\beta = 0$ [10]:

$$\dot{\mathbf{d}}^{t+\frac{1}{2}\Delta t} = \dot{\mathbf{d}}^{t} + \frac{1}{2}\Delta t\, \ddot{\mathbf{d}}^{t}\,; \tag{11a}$$

$$\mathbf{d}^{t+\Delta t} = \mathbf{d}^{t} + \Delta t\, \dot{\mathbf{d}}^{t+\frac{1}{2}\Delta t}\,; \tag{11b}$$

$$\ddot{\mathbf{d}}^{t+\Delta t} = \mathbf{M}^{-1}\left(\mathbf{f}^{\mathrm{ext},t+\Delta t} - \mathbf{f}^{\mathrm{int},t+\Delta t}\right)\,; \tag{11c}$$

$$\dot{\mathbf{d}}^{t+\Delta t} = \dot{\mathbf{d}}^{t+\frac{1}{2}\Delta t} + \frac{1}{2}\Delta t\, \ddot{\mathbf{d}}^{t+\Delta t}\,, \tag{11d}$$

In these equations Δt is the discrete time step and $\mathbf{d}$ represents the total array of degrees of freedom in the system: $\mathbf{d} = [\mathbf{a}, \mathbf{b}]^{\mathrm{T}}$.

For wave equations, the time step is limited by the speed of the stress wave in the material, or the dilatation speed c_{d}. In general, the simulation becomes unstable when the stress wave crosses a single element with a specific element length l_e within one time interval Δt. Hence, the critical interval is equal to:

$$\Delta t_{\mathrm{cr}} = \frac{l_e}{c_{\mathrm{d}}}\,. \tag{12}$$

Traditionally, the calculation time of an explicit time integration procedure is reduced significantly by using a diagonal mass representation $\mathbf{M}^{\mathrm{diag}}$ in equation (11c). In such a matrix, all terms are zero except for those on the diagonal. As a result, the system of equations (11c) is no longer coupled and each degree of freedom i can be solved separately:

$$d_i^{t+\Delta t} = \frac{f_i^{\mathrm{ext},t+\Delta t} - f_i^{\mathrm{int},t+\Delta t}}{M_i^{\mathrm{diag}}}\,. \tag{13}$$

A diagonal mass matrix can be constructed in several different ways. In the case of a standard continuum element formulation, it is possible to concentrate the mass of the element in its nodes. Each node then has a portion of the total mass of the element, which is placed on the diagonal term of the corresponding degrees of freedom in the mass matrix.

An alternative approach that is pursued here is to *lump* the mass matrix by replacing the diagonal term of the mass matrix by the total sum of the mass terms on the corresponding row. The off-diagonal terms are set to zero. In the case of a system with n unknowns, the diagonal mass matrix $\mathbf{M}^{\mathrm{diag}}$ can be written as:

$$M_i^{\mathrm{diag}} = \sum_{j}^{n} M_{ij}\,. \tag{14}$$

It can be seen from equation (9) that the mass matrix contains terms that couple the regular and additional degrees of freedom. By lumping the matrix, this information will be lost. In order to determine the effects of lumping, the different approaches are studied in the following problem, which is based on a numerical test by Xu and Needleman [11].

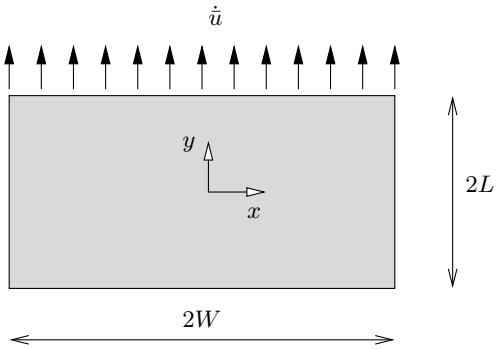


Fig. 2. Geometry and loading condition of the block.

4 Benchmark Problem

Consider the plane strain wave propagation block in Figure 2. The block has dimensions $L = 5\,\text{mm}$ and $W = 10\,\text{mm}$ and is made of PMMA with Young's modulus $E = 3.24\,\text{GPa}$, Poisson's ratio $\nu = 0.35$ and density $\rho = 1190\,\text{kg/m}^3$. The corresponding dilatational, shear and Rayleigh wave speeds for the given material are $c_\text{d} = 2090\,\text{m/s}$, $c_\text{s} = 1004\,\text{m/s}$ and $c_\text{R} = 938\,\text{m/s}$, respectively [12].

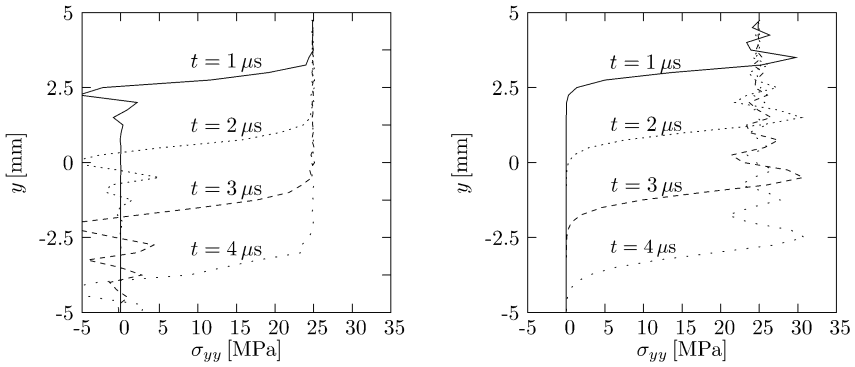
The block is not supported and is loaded by an impact velocity $\dot{\bar{u}}$, acting in the positive y direction on the top boundary of the block (at $y = +L$). The impact velocity is increased linearly to a constant value V within a certain rise time t_r, according to:

$$\dot{\bar{u}} = \begin{cases} Vt/t_r & \text{for} \quad t < t_r\,, \\ V & \text{for} \quad t \geq t_r\,. \end{cases} \tag{15}$$

In this experiment, the rise time is taken to be $t_r = 1.0 \cdot 10^{-7}\,\text{s}$ and the prescribed velocity $V = 10\,\text{m/s}$. The uni-axial stress wave created by these loading conditions will propagate through the block at approximately the dilatational wave speed, $c_\text{d} = 2090\,\text{m/s}$, carrying a tensile stress of 25.0 MPa.

Since the block and the loading conditions are symmetric about the y-axis, only the part in the positive x-axis is modelled using a regular mesh with 39×39 four node elements. The specific length of each element is $l_e = 0.256\,\text{mm}$. The displacements in the x-direction of the nodes at $x = 0.0\,\text{mm}$ are constrained in order to enforce symmetric boundary conditions. Regarding the specific length of an element and the magnitude of the dilatational wave speed, a sufficiently small time step of $\Delta t = 1.0 \cdot 10^{-9}\,\text{s}$ has been used in the simulations, which have been performed using both a lumped and a consistent representation of the mass matrix.

Figure 3 shows the profile of the σ_{yy} stress along the y-axis at $x = 0.0\,\text{mm}$ during the simulation. It can be seen that the average magnitude of the stress wave is in agreement with the analytical prediction of 25 MPa. The small



(a) Consistent mass matrix (b) Lumped mass matrix

Fig. 3. Stress curves σ_{yy} versus vertical position y measured along the centre of the plane strain wave specimen ($x = 0.0\,\mathrm{mm}$) at different times during the simulation.

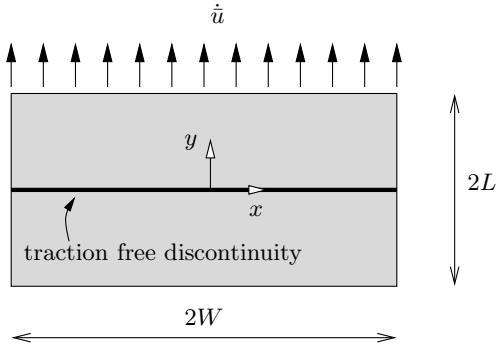
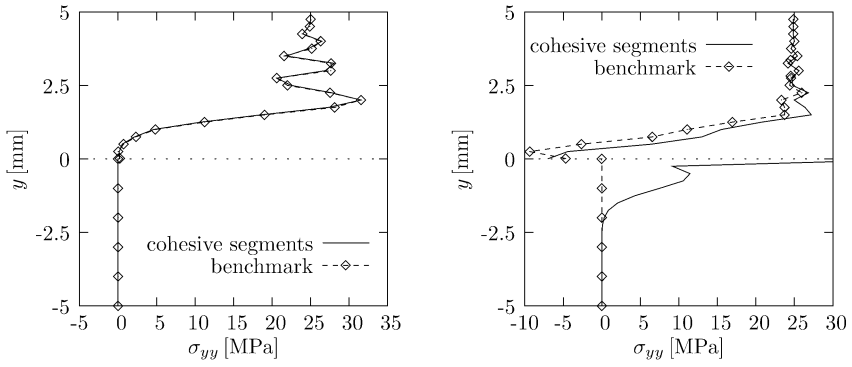


Fig. 4. Geometry and loading condition of the block with a traction free discontinuity.

stress oscillations of approximately 5.0 MPa can be contributed to the relatively coarse mesh that has been used in these simulations. The wave speed can be determined from the progression of the stress front with respect to time. Due to the numerical smoothing in the coarse mesh, the position of the stress wave cannot be determined exactly. Nevertheless, a rough estimation reveals that the wave speed is in good agreement with the analytically obtained dilatational speed $c_\mathrm{d} = 2090\,\mathrm{m/s}$.

A remarkable difference between the simulations with the lumped mass matrix and the simulations with the consistent mass matrix is the appearance of the stress overshoot. The lumped mass matrix gives rise to an extra overshoot of approximately 5.0 MPa on top of the wave front of 25.0 MPa, whereas in the case of a consistent mass matrix, the stress wave is preceded by a dip of 5.0 MPa.



(a) Consistent mass matrix (b) Lumped mass matrix

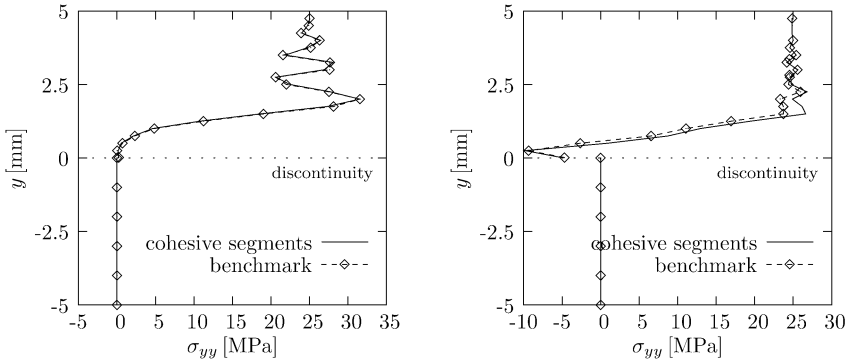
Fig. 5. Stress curves σ_{yy} versus vertical position y measured along the centre of the plane strain wave specimen ($x = 0.0\,\text{mm}$) at $t = 3\,\mu\text{s}$. The initial crack (denoted by the dotted line) is modelled as a traction free discontinuity which is constructed using a Heaviside jump function.

Subsequently, the block is divided into two parts by means of a horizontal crack at $y = 0\,\text{mm}$ crossing the entire width of the specimen, Figure 4. In the finite element mesh, this crack is represented as a traction free discontinuity. In this particular case, the discontinuity is represented using a standard Heaviside enhancement function, where $\mathcal{H}^+ = 1$ and $\mathcal{H}^- = 0$, see Table 1. The boundary conditions and material parameters of the specimen, as well as the impact velocity are not changed.

In the simulations, the stress wave carrying a tensile stress of 25 MPa will propagate from the top side of the specimen to reach the crack after approximately $t = 2\,\mu\text{s}$. Since the crack can be considered as a traction free boundary, the wave will reflect and move upwards. Obviously, the traction free crack prevents any kind of physical contact between the two parts of the specimen and the stresses in the lower part of the specimen will remain zero throughout the simulation.

Figure 5 shows the σ_{yy} stress curves along the y-axis in the centre of the specimen for both the mesh with a consistent and with a lumped representation of the mass matrix. In both graphs, the stress state at $t = 3\,\mu\text{s}$ is shown. At this instance, the stress wave has been reflected by the traction free crack. Both simulations are compared with a benchmark calculation in which the slit is created in the traditional way: by disconnecting adjacent elements. The benchmark calculations are performed using a consistent and a lumped mass matrix, respectively.

The simulation of the specimen with the crack using the consistent mass matrix shows perfect agreement with the benchmark simulation, see Figure 5 (a). The stress wave is properly reflected and, more importantly, the lower part of the specimen remains completely stress free. This is not the case in



(a) Consistent mass matrix (b) Lumped mass matrix

Fig. 6. Stress curves σ_{yy} versus vertical position y measured along the centre of the plane strain wave specimen ($x = 0.0\,\text{mm}$) at $t = 3\,\mu\text{s}$. The initial crack (denoted by the dotted line) is modelled as a traction free discontinuity which is constructed using a unit symmetric jump function.

the simulation with a lumped mass matrix, Figure 5 (b). Here, a considerable amount of the energy is transferred across the discontinuity, which results in a stress wave with an amplitude of 10 MPa in the lower part of the specimen. As a matter of fact, this spurious stress wave propagates with exactly the dilatational speed. Even more alarming is the extremely high stress peak of over 42 MPa in the integration points just below the discontinuity. Compared to this, the small mismatch of the stresses in the top part of the specimen, can be considered as a modest problem.

The simulations are now repeated with a different jump function. Instead of the Heaviside function a symmetric jump function is used, in which $\mathcal{H}^+ = 1$ and $\mathcal{H}^- = -1$, Table 1. The corresponding traction profiles are shown in Figure 6. Remarkably, the spurious stress wave reflections in the simulations with the lumped mass matrix representation have disappeared, Figure 6 (b). The stresses in the lower part of the specimen remain zero. Apparently, this can be contributed to the symmetric nature of the enhancement function. Nevertheless, the mismatch in the reflected stress wave in the top part of the specimen did not disappear, which reveals a small inconsistency in the current lumped matrix formulation. The stress waves in the simulation with the consistent mass matrix again show a perfect agreement to the benchmark simulations, see Figure 6 (a).

An alternative way to check the accuracy of the method is to look at the energy balance of the system. Obviously, the amount of external work that is used to deform the specimen must be equal to the total internal energy, which is composed of the strain energy and the kinetic energy. The external work $\mathcal{W}$ at time $t + \Delta t$ is obtained by integration in time, according to, see [13]:

$$\mathcal{W}^{t+\Delta t} = \mathcal{W}^t + \frac{1}{2}\Delta t \left(\dot{\mathbf{d}}^{t+\frac{1}{2}\Delta t}\right)^{\mathrm{T}} \left(\mathbf{f}^{\mathrm{ext},t} + \mathbf{f}^{\mathrm{ext},t+\Delta t}\right) . \tag{16}$$

When the specimen is loaded by a prescribed velocity instead of an external force, this relation can be written as:

$$\mathcal{W}^{t+\Delta t} = \mathcal{W}^t + \frac{1}{2}\Delta t \left(\dot{\mathbf{d}}_{\mathrm{p}}^{t+\frac{1}{2}\Delta t}\right)^{\mathrm{T}} \left(\mathbf{f}^{\mathrm{int},t} + \mathbf{f}^{\mathrm{int},t+\Delta t}\right) , \tag{17}$$

where $\dot{\mathbf{d}}_{\mathrm{p}}^{t+\frac{1}{2}\Delta t}$ denotes the prescribed velocity vector in which the unconstrained degrees of freedom are replaced by zeros. The internal strain energy is obtained as follows:

$$\mathcal{E}_{\mathrm{int}}^{t+\Delta t} = \mathcal{E}_{\mathrm{int}}^{t} + \frac{1}{2}\Delta t \left(\dot{\mathbf{d}}^{t+\frac{1}{2}\Delta t}\right)^{\mathrm{T}} \left(\mathbf{f}^{\mathrm{int},t} + \mathbf{f}^{\mathrm{int},t+\Delta t}\right) , \tag{18}$$

whereas the kinetic energy is given by:

$$\mathcal{E}_{\mathrm{kin}}^{t+\Delta t} = \frac{1}{2} \left(\dot{\mathbf{d}}^{t+\Delta t}\right)^{\mathrm{T}} \mathbf{M}\dot{\mathbf{d}}^{t+\Delta t} . \tag{19}$$

Energy conservation requires that the sum of internal and kinetic energy is equal to the total external work, or:

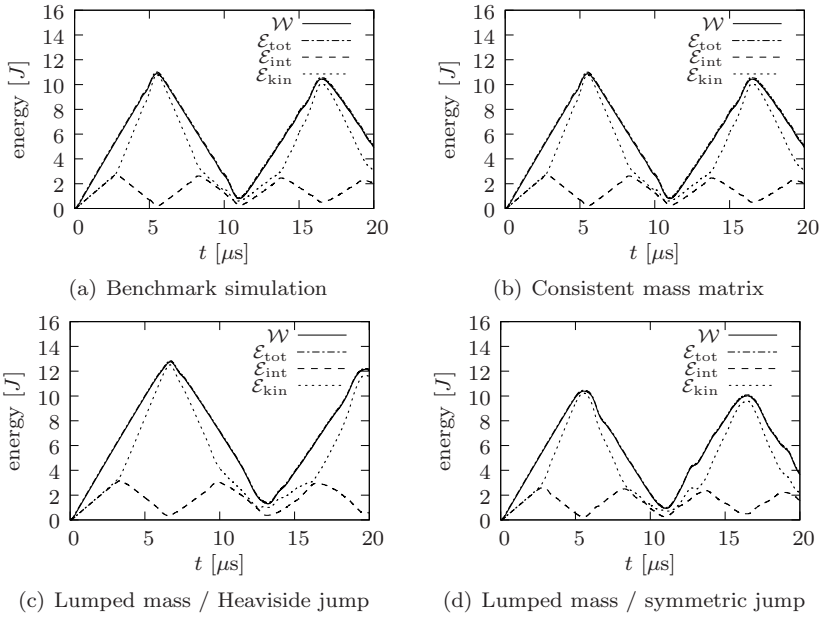
$$|\mathcal{E}_{\mathrm{kin}} + \mathcal{E}_{\mathrm{int}} - \mathcal{W}| < \varepsilon \, \max(|\mathcal{W}|, |\mathcal{E}_{\mathrm{int}}|, \mathcal{E}_{\mathrm{kin}}) , \tag{20}$$

where ε is a small tolerance. When this tolerance is on the order of 10^{-2} or smaller at each time increment, it can be concluded that energy is conserved [13].

The energy variations as a function of time for the simulations with the traction free discontinuity are shown in Figure 7. Figure 7 (a) shows the variation of energy for the benchmark calculation in which the crack is modelled in the traditional way by disconnecting adjacent elements. In this particular case, a consistent mass matrix is used. The energy variations for the calculation with a lumped mass matrix are nearly identical. It can be seen that in these calculations, the kinetic and internal strain energy sum up to the external work, which indicates that energy is conserved. Note that the line that represents the total internal energy $\mathcal{E}_{\mathrm{tot}}$ is hardly visible in Figures 7 (a) to (d) since it is overlapped by the line representing the external work $\mathcal{W}$.

Figures 7 (b) to (d) show the energy variations for the simulations in which the slit is represented by a traction free discontinuity. In (b) a consistent mass matrix has been used, in combination with a Heaviside jump function for the discontinuity. In the calculation depicted in (c) a lumped mass matrix has been used. Here, the discontinuity has been modelled with a Heaviside step function. Finally, Figure 7 (d) shows the energy variation of the simulation with a lumped mass matrix in combination with a unit symmetric step function for the discontinuity.

Although in both simulations with the lumped mass matrix energy is conserved, the variation of total internal energy in time is different than the



(a) Benchmark simulation

(b) Consistent mass matrix

(c) Lumped mass / Heaviside jump

(d) Lumped mass / symmetric jump

Fig. 7. Energy variations as a function of time. Figure (a) shows the benchmark simulation with a traditional finite element mesh. The other figures show the calculations with a traction free discontinuity with (b) a consistent mass matrix, (c) a lumped mass matrix and a Heaviside jump function and (d) a lumped mass matrix and a symmetric jump function.

variation of total energy in the benchmark case. For the calculations with the Heaviside jump this was to be expected. It was demonstrated before that a significant amount of energy was transferred across the discontinuity. As a result, the magnitude of the stress waves reflecting back from the discontinuity is smaller, which in turn gives rise to a different behaviour at the clamped edge. The deviation of the total energy in the case of a lumped mass matrix in combination with a unit symmetric jump function, Figure 7 (d) may come as a surprise. In the previous tests, this combination did not give rise to spurious stress wave reflections into the lower part of the specimen. Nevertheless, the magnitude of the σ_{yy} stress carried by the reflecting stress wave is slightly different than the magnitude of this stress in the benchmark simulation, see Figure 6 (b). It appears that this relatively small deviation gives rise to a difference in the total energy variation.

Despite the differences in energy variations, in all cases, the total energy is well conserved. Apparently, the traction free discontinuity does not give rise to the creation or dissipation of internal energy, even when a lumped mass matrix and a Heaviside jump function are used. The magnitudes of the tolerance ε for all four cases are shown in Figure 8. Apart from a few peaks, the average value of the tolerance is indeed on the order of 10^{-2}. The large peak at the start of

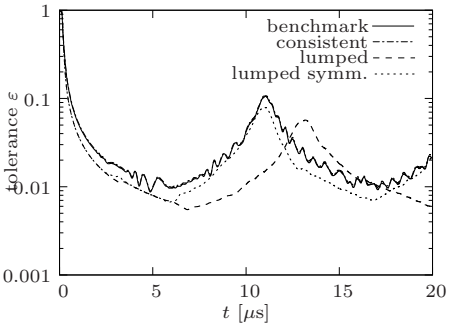


Fig. 8. Energy tolerances as a function of time for the four simulations: the benchmark simulation, the traction free discontinuity with a consistent mass matrix, the lumped mass matrix in combination with the Heaviside jump function and the lumped mass matrix in combination with the symmetric jump function.

the simulation is due to the fact that at this instant the prescribed velocity $\bar{u}$ is increased to its final value V in a relatively short amount of time. The second peak after approximately $t = 10\,\mu$s is caused by the reflecting stress wave that reaches the crack of the specimen for the second time. This marks a sudden change in the direction of stress wave propagation, which results in an abrupt change of external work and therefore in an increase of the energy mismatch. In general, these boundary effects are less significant when a denser finite element mesh is used in which the amount of degrees of freedom that are constrained is relatively small.

In general, the simulations in which a consistent mass matrix is used give results that are identical to the benchmark simulations. When a lumped mass matrix is used, different jump functions to give rise to different results.

5 Stability of the Time-Integration Algorithm

It has already been observed by Babuška and Melenk [9] that in most cases the condition of the Jacobian decreases significantly when base functions are added to shape functions by using the partition of unity method. It was demonstrated by Wells and Sluys [2] that the stiffness matrix in implicit calculations becomes ill-conditioned when a discontinuity crosses an element close to a node. This anomaly however can be solved by enhancing only those nodes that have a significant contribution to the stiffness matrix.

In explicit simulations a more or less equivalent problem occurs. It appears that when an element is crossed by a discontinuity, the two separated parts can be considered as individual elements, each with a smaller effective length l_e than the original element. In principle, an element can be crossed by a discontinuity in such a way that one of the two resulting portions of the element becomes so small that, according to relation (12), the critical time

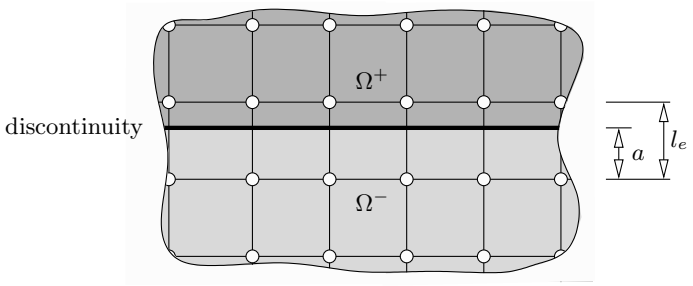


Fig. 9. Detail of the quadrilateral finite element mesh that has been used to determine the critical time step for various positions of a discontinuity.

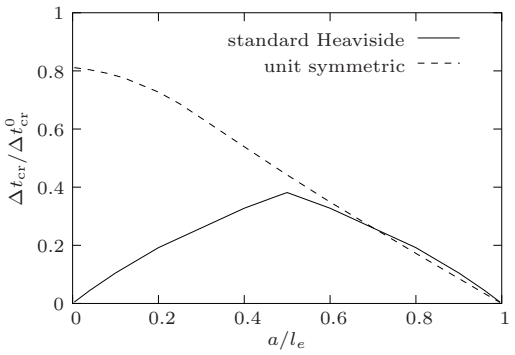


Fig. 10. The critical time step vs. the position of the discontinuity within the element. s denotes the distance of the discontinuity with respect to the nodes in the Ω^- part of the domain.

increment for a stable solution procedure will almost be infinitesimal. In such a situation, numerical solutions, such as the selective enhancement of nodes, do not work.

In order to determine the exact effect of the position of the discontinuity on the critical time step of the explicit time integration method, the simulation of the plane strain wave specimen with an initial slit is considered again. In this case, the specimen consists of 40×40 quadrilateral elements with an specific element length $l_e = 0.25\,\text{mm}$. The position of the horizontal traction free discontinuity that represents the initial slit is varied between $a = 0.0\,\text{mm}$ and $a = 0.25\,\text{mm}$ so that it covers all positions within a single element, see Figure 9. The material parameters and the boundary and loading conditions remain the same. The simulations are performed using a standard Heaviside function and a unit-symmetric enhancement function.

The critical time increment Δt_{cr} is shown as a function of the position of the discontinuity within the element in Figure 10. The critical time increments are normalised by the critical time step in case of a simulation without a

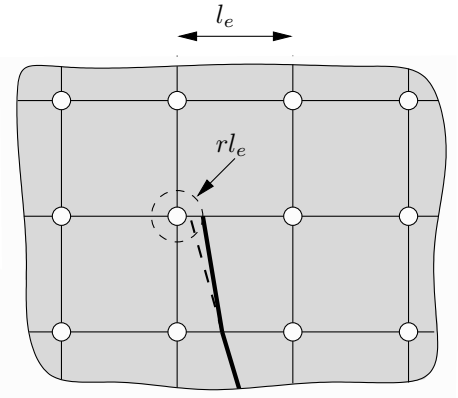


Fig. 11. Overview of the algorithm that prevents a discontinuity to cross an element boundary in the vicinity of a supporting node. The grey dashed line denotes the original position of the discontinuity. Since it crosses the node within a distance rl_e (denoted by the dashed circle), it is moved away from the node.

discontinuity Δt^0_{cr}. This value is found to be $0.1172\,\mu$s which is relatively close to the estimation from equation (12): $\Delta t^0_{\text{cr}} = 0.1196\,\mu$s.

As expected, the critical time increment decreases almost linearly to zero when the discontinuity approaches the boundary of the element. In the case of the standard Heaviside enhancement function, there is no difference whether this boundary is in the Ω^+ or the Ω^- part of the domain. When using the unit-symmetric enhancement function, this effect only occurs when the discontinuity is in the vicinity of nodes in the Ω^+ part of the domain. On the other side of the element, the critical time increment increases up to over 80% of the nominal critical increment when the discontinuity approaches the boundary in the Ω^- part of the domain.

The problem of infinitesimal critical time increments can be avoided by using an alternative lumping technique, as demonstrated by Menouillard et al. [14]. Unfortunately, this lumping procedure suffers from the spurious transfer of stress waves as reported in section 4 and will not be considered in the scope of this paper.

In the case of the simulation of propagating cracks, infinitesimal critical increments are only avoided when the discontinuity is not allowed to cross an element boundary within a certain distance rl_e to a node, where r is an offset factor between 0.0 and 0.5, see Figure 11 [8]. When r is set to 0.0, the trajectory of the discontinuity is not confined at all. When $r = 0.5$ the discontinuity is only allowed to cross an element boundary exactly through its centre point, which leaves as many possible crack trajectories as the original approach by Xu and Needleman [11]. Intermediate values of r will lead to partially modified crack trajectories.

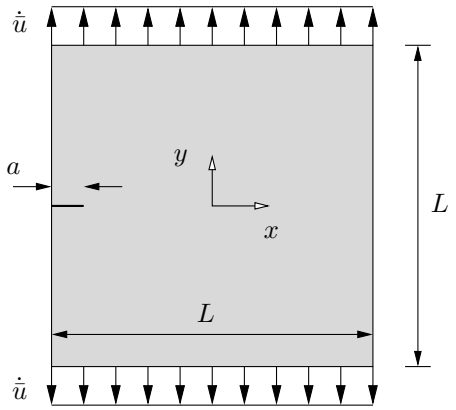


Fig. 12. Geometry and loading conditions of a square block with an initial side crack.

Because of the nearly linear relation between the critical time step and the distance between the discontinuity, the stability requirement can be expanded as follows:

$$\Delta t = \alpha \frac{r\, l_e}{c_{\mathrm{d}}}\,, \tag{21}$$

where α is an additional factor to compensate for shock wave effects at crack propagation [11], which is normally set to 0.1. In the simulations in the next section, r is taken to be 0.1, which implies that the crack is deflected with at most a few degrees.

6 A Propagating Discontinuity

In the previous section, it was observed that the inclusion of a discontinuity in a finite element formulation can influence the robustness of the explicit time integration scheme. In the following test, the effects of adding new degrees of freedom to the system in order to accommodate a propagating discontinuity is considered.

Consider the square block made of PMMA with length sides $L = 3\,\mathrm{mm}$ as shown in Figure 12. The block contains an edge crack that penetrates $a = 0.25\,\mathrm{mm}$ into the material. The block is loaded in tension by two pulse loads that are applied to the top and bottom edges respectively. In the finite element simulations, these pulses are represented by prescribed velocities with magnitudes $V = 6\,\mathrm{m/s}$. This magnitude is reached after a rise time $t_r = 1.0{\cdot}10^{-7}\,\mathrm{s}$, see equation (15).

Because of symmetry of both the geometry of the specimen as well as the boundary conditions, it is assumed that the crack will propagate in a straight line and will open in a pure tensile mode. In these calculations, the

discontinuity is extended along the x-axis when the σ_{yy} stress in the element ahead of the tip of the discontinuity exceeds the strength of the material $t_{\max}$. The discontinuity is extended until it touches the boundary of the next element. All nodes of the element are hereby enhanced apart from the ones that support the terminus of the discontinuity to ensure a zero opening at the tip [2].

The opening of the discontinuity is governed by a cohesive constitutive relation [2, 8]. The normal component of the traction t_n is related to the normal component of the displacement jump v_n, according to [15]:

$$t_{\mathrm{n}} = t_{\max}\left(1 - \frac{v_{\mathrm{n}}}{v_{\mathrm{n,cr}}}\right) . \tag{22}$$

where $v_{\mathrm{n,cr}}$ is the characteristic length of the cohesive law at which the crack has fully developed and the tractions have reduced to zero. This parameter is related to the fracture toughness $\mathcal{G}_c$, or the area under the softening curve and can be determined as follows:

$$v_{\mathrm{n,cr}} = \frac{2\mathcal{G}_c}{t_{\max}} . \tag{23}$$

In this simulation, the strength of the material is set to $t_{\max} = 1.0{\cdot}10^8\,\mathrm{N/m^2}$ and the fracture toughness $\mathcal{G}_c = 700\,\mathrm{N/m}$. Since the traction at the discontinuity in normal direction at zero opening is equal to the strength of the material, traction continuity is guaranteed [16].

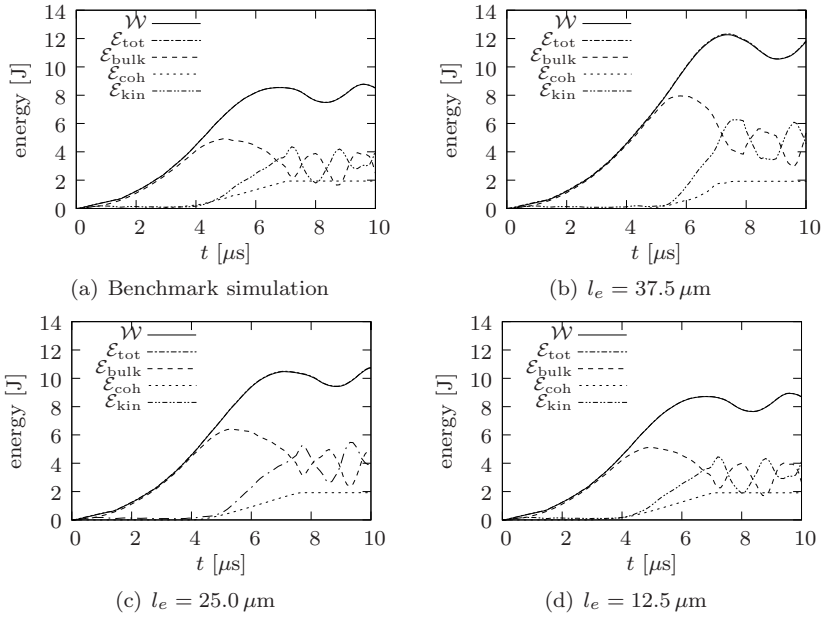
The block is analysed with three different finite element meshes that consist of quadrilateral elements only. In the first mesh, the specific length of an element is $l_e = 37.5\,\mu\mathrm{m}$, in the second mesh this length is $l_e = 25.0\,\mu\mathrm{m}$ and in the third mesh $l_e = 12.5\,\mu\mathrm{m}$. Considering equation (21) a timestep $\Delta t = 1.0 \cdot 10^{-9}\,\mathrm{s}$ is used in all simulations.

Since the trajectory of the crack is known beforehand, a comparative study is performed in which the crack is represented by a weak interface. In this benchmark calculation, a discontinuity is inserted over the full width of the specimen at the start of the analysis. The first 0.25 mm of this segment are traction free, the remaining part is equipped with Xu-Needleman's cohesive law [11], with the same cohesive properties as mentioned above. Apart from the fact that the cohesive zone in this case has an initial non-zero compliance, the overall behaviour of this model must be nearly identical. However, note that in this case, no degrees of freedom are added during the simulation.

The energy variations of the simulations are given in Figure 13. In these plots, a distinction is made between the contribution of the bulk material and the cohesive zone to the internal energy $\mathcal{E}_{\mathrm{int}}$:

$$\mathcal{E}_{\mathrm{int}} = \mathcal{E}_{\mathrm{bulk}} + \mathcal{E}_{\mathrm{coh}} , \tag{24}$$

where $\mathcal{E}_{\mathrm{bulk}}$ is the internal energy of the bulk material, which is defined as:



(a) Benchmark simulation (b) $l_e = 37.5\,\mu\text{m}$

(c) $l_e = 25.0\,\mu\text{m}$ (d) $l_e = 12.5\,\mu\text{m}$

Fig. 13. Energy variations as a function of time. Figure (a) shows the benchmark simulation with an initial discontinuity with Xu-Needleman's cohesive relation. The other figures show the calculations with a propagating discontinuity in meshes with different specific element lengths (b) 37.5 μm, (c) 25.0 μm and (d) 12.5 μm.

$$\mathcal{E}_{\text{bulk}}^{t+\Delta t} = \mathcal{E}_{\text{bulk}}^{t} + \frac{1}{2}\Delta t \left(\dot{\mathbf{d}}^{t+\frac{1}{2}\Delta t}\right)^{\text{T}} \left(\mathbf{f}_{\text{bulk}}^{\text{int},t} + \mathbf{f}_{\text{bulk}}^{\text{int},t+\Delta t}\right) , \tag{25}$$

where $\mathbf{f}_{\text{bulk}}^{\text{int}}$ is the contribution of the *bulk material* to the internal forces, see equation (9). For the cohesive zone, a similar relation can be formulated:

$$\mathcal{E}_{\text{coh}}^{t+\Delta t} = \mathcal{E}_{\text{coh}}^{t} + \frac{1}{2}\Delta t \left(\dot{\mathbf{d}}^{t+\frac{1}{2}\Delta t}\right)^{\text{T}} \left(\mathbf{f}_{\text{coh}}^{\text{int},t} + \mathbf{f}_{\text{coh}}^{\text{int},t+\Delta t}\right) , \tag{26}$$

where $\mathbf{f}_{\text{coh}}^{\text{int}}$ is the contribution of the *cohesive zone* to the internal forces, see (9).

Figure 13 (a) shows the variations of the total work $\mathcal{W}$, the sum of internal energies $\mathcal{E}_{\text{tot}}$ and the variations of the individual terms $\mathcal{E}_{\text{bulk}}$, $\mathcal{E}_{\text{coh}}$ and $\mathcal{E}_{\text{kin}}$ for the benchmark simulation with a predefined discontinuity in combination with Xu-Needleman's cohesive relation, i.e. the simulation in which no degrees of freedom are added during the simulation. In this particular figure, the energy variations of the simulation with the dense mesh are shown, but the results with the coarser meshes are nearly identical. Figures 13 (b) to (d) show the variations of the simulations with the coarse, medium and dense mesh respectively. It appears that in all four cases, the total internal energy is in good agreement with the external work of the simulations. This can also be

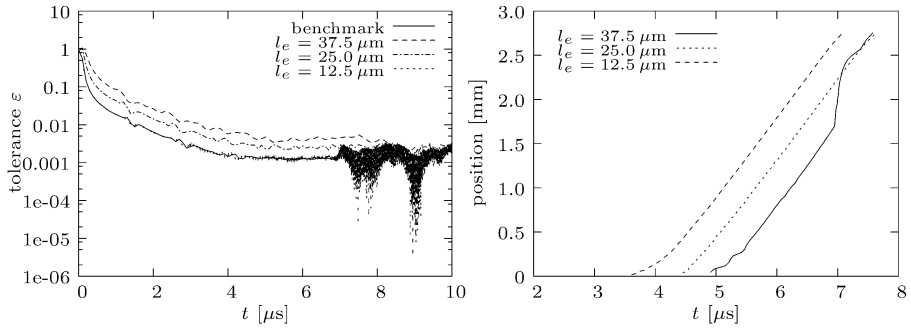


Fig. 14. (a) Energy tolerances as a function of time for the four simulations of a propagating straight crack through a square block. (b) Position of the crack tip as a function of time for the four simulations of a propagating straight crack through a square block.

seen in Figure 14 (a) in which the error tolerances of the four simulations are plotted. In none of the cases ε exceeds the value of 0.01. It can be concluded that irrespective of the coarseness of the finite element mesh, the creation of new degrees of freedom does not destroy the energy balance. This is a result of the fact that in the cohesive constitutive relation, the value of the strength $t_{\mathrm{n},0}$ is taken to be equal to the the stress state that caused the discontinuity to propagate. Effects of overshoot are thus canceled and stress continuity is ensured [16]. The oscillations at the end of the simulations are a result of the fact that at the instant of the extension of the discontinuity, the values of the internal energy are not uniquely defined. The displacement and internal force terms in equations (25) and (26) that correspond to the degrees of freedoms that have been added at time $t + \Delta t$ are not defined at time t, which gives rise to a small error.

Nevertheless, a remarkable difference in the simulations with the coarser meshes can be observed, Figure 13 (b) and (c). Here, the total work is significantly larger than the total work in the benchmark simulation (a) and the simulation with the dense mesh (d). This a consequence of the fact that the *averaged* σ_{yy} stress in the element ahead of the tip is used to extend the discontinuity. Since the magnitude of this stress is smaller than the actual stress at the tip, the new segment is extended too late, particularly in the simulations with a coarse mesh. This allows the bulk material to store too much internal energy prior to crack propagation. This can also be concluded from Figure 14 (b) where the position of the crack tip are plotted as a function of time. Here, the position of the crack tip is defined as the first point along the discontinuity where the tractions have reduced to 5% the value of the strength of the material.

The coarseness of the mesh does not have an effect on the internal energy dissipated by the cohesive segment. In all simulations with a propagating

segment, the final value of the cohesive internal energy $\mathcal{E}_{\text{coh}}$, when the crack has completely crossed the specimen is equal to $\mathcal{E}_{\text{coh}} = 1.92\,\text{J}$, which is exactly identical to the length of the crack, 2.75 mm times the fracture toughness of the material $\mathcal{G}_c = 700\,\text{N/m}$. In the benchmark simulation, with an initial cohesive segment, this value is slightly bigger, $\mathcal{E}_{\text{coh}} = 1.93\,\text{J}$, which is probably due to the fact that prior to cracking, the cohesive zone, with its initial non-zero compliance, already contributes to the internal energy.

7 Conclusions

In this paper, the accuracy of partition of unity based discontinuous elements in explicit simulations has been discussed. In the first simulation, the consistency of the method was checked by means of a benchmark test in which a solid block is divided in two parts by a slit over the full width of the specimen. In the finite element simulations, the slit is represented by a traction free discontinuity. When the mass matrix is lumped, spurious transfer of stress waves across the discontinuity can be observed. This anomaly disappears when the Heaviside function is replaced by a unit-symmetric jump function. However, from an energetic point of view, the numerical results still showed small deviations. Only use of the consistent mass matrix guarantees correct results. This requires the solution of a fully coupled system. However, the additional computational effort can be minimised by employing the observation that the mass matrix only changes when new degrees of freedom are added, which normally occurs in a limited number of steps. The decomposition of the matrix, which is by far the most expensive part of the solution process, only needs to be preformed after these steps.

In a second benchmark problem, the stability of the solution procedure has been investigated. In standard finite element implementations, the critical step size is approximately equal to the dilatational wave speed of the material divided by the element specific length. In the case of discontinuous elements, the specific length appears to be equal to the specific length of the smallest part of a discontinuous element. Depending on the position of the discontinuity, this length can go to zero, which results in an infinitesimal time step. The only remedy to this problem is to prevent discontinuities from crossing an element border close to a node.

In the final problem, the effect of adding new degrees of freedom for a propagating discontinuity is studied. The behaviour of the fracture process is represented by a cohesive zone model, which is known to be energetically consistent. It has been checked that the energy is conserved, regardless of the mesh size of the specimen. An important requirement however is that traction continuity is guaranteed when the discontinuity is extended.

References

1. Belytschko T and Black T (1999) International Journal for Numerical Methods in Engineering 45 (5): 601–620
2. Wells G N and Sluys L J (2001) International Journal for Numerical Methods in Engineering 50 (12): 2667–2682
3. Moës N and Belytschko T (2002) Engineering Fracture Mechanics 69: 813–833
4. Wells G N, de Borst R and Sluys L J (2002) International Journal for Numerical Methods in Engineering 54 (9): 1333–1355
5. Remmers J J C, Wells G N and de Borst R (2003) International Journal for Numerical Methods in Engineering 58 (13): 2013–2040
6. Belytschko T, Chen H, Xu J and Zi G (2003) International Journal for Numerical Methods in Engineering, 58: 1873–1905
7. Réthoré J, Gravouil A and Combescure A (2005) International Journal for Numerical Methods in Engineering 63 (5): 631–659
8. Remmers J J C, de Borst R and Needleman A (2007) submitted to Journal of the Mechanics and Physics of Solids
9. Babuška I and Melenk J M (1997) International Journal for Numerical Methods in Engineering 40 (4): 727–758
10. Belytschko T, Chiapetta R L and Bartel H D (1976) International Journal for Numerical Methods in Engineering 10 (3): 579–596
11. Xu X P, Needleman A (1994) Journal of the Mechanics and Physics of Solids 42: 1397–1434
12. Freund L B (1998) Dynamic Fracture Mechanics, Cambridge University Press, Cambridge
13. Belytschko T, Liu W K and Moran B (2000) Nonlinear Finite Elements for Continua and Structures. John Wiley and Sons Ltd., England
14. Menouillard T, Réthoré J, Combescure A and Bung H (2006) International Journal for Numerical Methods in Engineering 68 (9): 911–939
15. Camacho G T and Ortiz M (1996) International Journal of Solids and Structures 33: 2899–2938
16. Papoulia K D, Sam S H and Vavasis S A (2003) International Journal for Numerical Methods in Engineering 58 (5): 679–701

A discrete model for the propagation of discontinuities in a fluid-saturated medium

Julien Réthoré[1], René de Borst[12], and Marie-Angèle Abellan[3]

[1] Faculty of Aerospace Engineering, Delft University of Technology, Delft, Netherlands r.deborst@tudelft.nl
[2] LaMCoS - UMR CNRS 5514, INSA de Lyon, Villeurbanne, France
[3] LTDS-ENISE - UMR CNRS 5513, Saint-Etienne, France

Summary. The first part of this manuscript discusses a finite element method that captures arbitrary discontinuities in a two-phase medium by exploiting the partition-of-unity property of finite element shape functions. The fluid flow away from the discontinuity is modelled in a standard fashion using Darcy's relation, and at the discontinuity a discrete analogy of Darcy's relation is used. Subsequently, dynamic shear banding is studied numerically for a biaxial, plane-strain specimen. A Tresca-like as well as a Coulomb criterion are used as nucleation criterion. Decohesion is controlled by a mode-II fracture energy, while for the Coulomb criterion, frictional forces are transmitted across the interface in addition to the cohesive shear tractions. The effect of the different interface relations on the onset of cavitation is studied.

Key words: shear band, dynamic fracture, two-phase medium, partition-of-unity method

1 Introduction

Broadly speaking, two approaches exist for the numerical analysis of the nucleation and propagation of discontinuities in solids, such as cracks, shear bands and faults. Within the classical theory of continuum mechanics, the approach in which discontinuities are distributed over a finite volume, so that relative displacements across the faces of a discontinuity are transformed into strains, is perhaps the most natural. Also in an engineering sense, it offers advantages, since there is no need to keep track of all individual microcracks that arise in a solid. Indeed, for computations of large structures, any attempt to model each individual crack would exceed even the currently available computing power. However, from a theoretical point of view, the modelling of discontinuities in a distributed or smeared sense has a limitation, namely that at a certain level of accumulated damage, the set of governing equations locally changes

Alain Combescure et al. (eds.), IUTAM Symposium on Discretization Methods for Evolving Discontinuities, 323–342.

character, from elliptic to hyperbolic for quasi–static loadings, and from hyperbolic to elliptic for dynamic loadings. Unless a regularisation is applied, the resulting initial/boundary value problem becomes ill–posed, resulting in numerical solutions that depend severely on the discretisation [1].

Intuitively the most appealing approach is to model discontinuities in a discrete manner, thus reflecting the change in topology that actually takes place in the solid when a discontinuity propagates. Recently, such finite element methods have been constructed that exploit the partition-of-unity property of finite element shape functions [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]. While preserving the original discretisation, the addition of extra degrees-of-freedom to nodes whose support is crossed by a discontinuity allows to construct two continuous displacement fields that are separated by a Heaviside function at the discontinuity. As a consequence, discontinuities can propagate, not biased by the original discretisation.

Many problems in geomechanics involve the coupling of the set of equations that describe the stress evolution and those which describe diffusion-type processes, e.g. water or ion transport. Indeed, hydro-mechanical interactions have been recognised to play a crucial role in geotechnical, petroleum and mining engineering since the pioneering works by Terzaghi [13] and Biot [14]. It is the purpose of this manuscript to formulate a numerical model that is capable of describing dynamic shear band propagation in a porous medium, with a solid skeleton and an interstitial fluid as the constituent phases, in a discrete, mesh-independent manner. The model exploits the partition-of-unity property of finite element shape functions, and can therefore be considered to be an extension to earlier works on fracture for single-phase media. On the other hand, the present methodology can be extended in a fairly straightforward manner to introduce discontinuities, including cracks, in initial value problems where several diffusion-type problems play a role.

The manuscript starts with a concise derivation of the balance equations for a fluid-saturated porous medium. Subsequently, the general methodology and the assumptions regarding the introduction of a discontinuity in a finite element model are discussed. The approach is specialised to a medium where the (discrete) failure mode is caused by exhaustion of the shear stress capacity on a critical plane. Next, studies are carried out for a plane-strain, biaxial specimen. Two different nucleation criteria for shear band propagation are investigated, and the role of localisation on cavitation in a fluid-saturated porous medium is highlighted.

2 Balance equations

We consider a two-phase medium subject to the restriction of small displacement gradients and small variations in the concentrations [15]. Furthermore, the assumptions are made that there is no mass transfer between the constituents and that the processes which we consider, occur isothermally. With

these assumptions, the balances of linear momentum for the solid and the fluid phases read:

$$\nabla \cdot \boldsymbol{\sigma}_\pi + \hat{\mathbf{p}}_\pi + \rho_\pi \mathbf{g} = \frac{\partial(\rho_\pi \ \mathbf{v}_\pi)}{\partial t} + \nabla \cdot (\rho_\pi \mathbf{v}_\pi \otimes \mathbf{v}_\pi) \tag{1}$$

with $\boldsymbol{\sigma}_\pi$ the stress tensor, ρ_π the apparent mass density, and $\mathbf{v}_\pi$ the absolute velocity of constituent π. As in the remainder of this paper, $\pi = s, f$, with s and f denoting the solid and fluid phases, respectively. Further, $\mathbf{g}$ is the gravity acceleration and $\hat{\mathbf{p}}_\pi$ is the source of momentum for constituent π from the other constituent, which takes into account the possible local drag interaction between the solid and the fluid. Evidently, the latter source terms must satisfy the momentum production constraint:

$$\sum_{\pi=s,f} \hat{\mathbf{p}}_\pi = \mathbf{0} \tag{2}$$

We now neglect convective terms and the gravity acceleration, so that the momentum balances reduce to:

$$\nabla \cdot \boldsymbol{\sigma}_\pi + \hat{\mathbf{p}}_\pi = \rho_\pi \frac{\partial \mathbf{v}_\pi}{\partial t} \tag{3}$$

Adding both momentum balances, and taking into account eq. (2), one obtains the momentum balance for the mixture:

$$\nabla \cdot \boldsymbol{\sigma} - \rho_s \frac{\partial \mathbf{v}_s}{\partial t} - \rho_f \frac{\partial \mathbf{v}_f}{\partial t} = \mathbf{0} \tag{4}$$

where the stress is, as usual, composed of a solid and a fluid part,

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}_s + \boldsymbol{\sigma}_f \tag{5}$$

For relatively slow dynamic loadings, the assumption is often made that the accelerations of the solid and of the fluid are equal: $\frac{\partial \mathbf{v}_s}{\partial t} \approx \frac{\partial \mathbf{v}_f}{\partial t}$. With the mass density of the mixture, $\rho = \rho_s + \rho_f$, the balance of momentum (4) reduces to:

$$\nabla \cdot \boldsymbol{\sigma} - \rho \frac{\partial \mathbf{v}_s}{\partial t} = \mathbf{0} \tag{6}$$

Numerical analyses are usually conducted with the latter equation as balance of momentum, cf [16], but the accuracy of this assumption is seldom quantified.

In a similar fashion as for the balances of momentum, one can write the balance of mass for each phase as:

$$\frac{\partial \rho_\pi}{\partial t} + \nabla \cdot (\rho_\pi \mathbf{v}_\pi) = 0 \tag{7}$$

Again neglecting convective terms, the mass balances can be simplified to give:

$$\frac{\partial \rho_\pi}{\partial t} + \rho_\pi \nabla \cdot \mathbf{v}_\pi = 0 \tag{8}$$

We multiply the mass balance for each constituent π by its volumetric ratio n_π, add them and utilise the constraint

$$\sum_{\pi=s,f} n_\pi = 1 \tag{9}$$

to give:

$$\nabla \cdot \mathbf{v}_s + n_f \nabla \cdot (\mathbf{v}_f - \mathbf{v}_s) + \frac{n_s}{\rho_s}\frac{\partial \rho_s}{\partial t} + \frac{n_f}{\rho_f}\frac{\partial \rho_f}{\partial t} = 0 \tag{10}$$

The change in the mass density of the solid material is related to its volume change by:

$$\nabla \cdot \mathbf{v}_s = -\frac{K_s}{K_t}\ \frac{n_s}{\rho_s}\frac{\partial \rho_s}{\partial t} \tag{11}$$

with K_s the bulk modulus of the solid material and K_t the overall bulk modulus of the porous medium. Using the definition of the Biot coefficient, $1-\alpha = K_t/K_s$ [16], this equation can be rewritten as

$$(\alpha - 1)\nabla \cdot \mathbf{v}_s = \frac{n_s}{\rho_s}\frac{\partial \rho_s}{\partial t} \tag{12}$$

For the fluid phase, a phenomenological relation is assumed between the incremental changes of the apparent fluid mass density and of the fluid pressure p [16]:

$$\frac{1}{Q}\mathrm{d}p = \frac{n_f}{\rho_f}\mathrm{d}\rho_f \tag{13}$$

with the overall compressibility, or Biot modulus

$$\frac{1}{Q} = \frac{\alpha - n_f}{K_s} + \frac{n_f}{K_f} \tag{14}$$

where K_f is the bulk modulus of the fluid. Inserting relations (12) and (13) into the balance of mass of the total medium, eq. (10), gives:

$$\alpha \nabla \cdot \mathbf{v}_s + n_f \nabla \cdot (\mathbf{v}_f - \mathbf{v}_s) + \frac{1}{Q}\frac{\partial p}{\partial t} = 0 \tag{15}$$

The field equations, i.e. the balance of momentum of the saturated medium, eq. (4), and the balance of mass, eq. (15), are complemented by the boundary conditions

$$\mathbf{n}_\Gamma \cdot \boldsymbol{\sigma} = \mathbf{t}_p \quad , \quad \mathbf{v} = \mathbf{v}_p \tag{16}$$

which hold on complementary parts of the boundary $\partial\Omega_t$ and $\partial\Omega_v$, with $\Gamma = \partial\Omega = \partial\Omega_t \cup \partial\Omega_v$, $\partial\Omega_t \cap \partial\Omega_v = \emptyset$, $\mathbf{t}_p$ being the prescribed external traction and $\mathbf{v}_p$ the prescribed velocity, and

$$n_f(\mathbf{v}_f - \mathbf{v}_s) \cdot \mathbf{n}_\Gamma = q_p \quad , \quad p = p_p \tag{17}$$

which hold on complementary parts of the boundary $\partial\Omega_q$ and $\partial\Omega_p$, with $\Gamma = \partial\Omega = \partial\Omega_q \cup \partial\Omega_p$ and $\partial\Omega_q \cap \partial\Omega_p = \emptyset$, q_p and p_p being the prescribed outflow of pore fluid and the prescribed pressure, respectively. The initial conditions which specify the displacements $\mathbf{u}_\pi$, the velocities $\mathbf{v}_\pi$, and the pressure field at $t = 0$:

$$\mathbf{u}_\pi(\mathbf{x}, 0) = \mathbf{u}_\pi^0, \quad \mathbf{v}_\pi(\mathbf{x}, 0) = \mathbf{v}_\pi^0, \quad p(\mathbf{x}, 0) = p^0 \tag{18}$$

close the initial value problem.

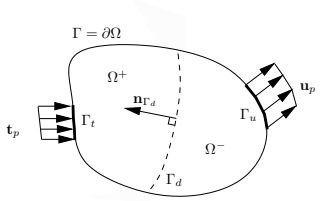


Fig. 1. Body composed of continuous displacement fields at each side of the discontinuity Γ_d

3 Discontinuities in a two-phase medium

A finite element method that can accommodate the propagation of discontinuities through elements was proposed by Belytschko and co-workers [3, 4], exploiting the partition-of-unity property of finite element shape functions [2]. Since finite element shape functions φ_j form partitions of unity, $\sum_{j=1}^{n} \varphi_j = 1$ with n the number of nodal points, the components v_i of a velocity field $\mathbf{v}$ can be interpolated as

$$v_i = \sum_{j=1}^{n} \varphi_j \left(\dot{\bar{a}}_j + \sum_{k=1}^{m} \psi_k \dot{\tilde{a}}_{jk} \right) \tag{19}$$

with $\bar{a}_j$ the 'regular' nodal degrees-of-freedom for the displacements, ψ_k the enhanced basis terms, and $\tilde{a}_{jk}$ the additional displacement degrees-of-freedom at node j which represent the amplitude of the kth enhanced basis term ψ_k. Next, we consider a domain Ω that is crossed by a single discontinuity at Γ_d (see Figure 1). The velocity field $\mathbf{v}$ can be written as the sum of two continuous velocity fields $\bar{\mathbf{v}}$ and $\tilde{\mathbf{v}}$:

$$\mathbf{v} = \bar{\mathbf{v}} + \mathcal{H}_{\Gamma_d} \tilde{\mathbf{v}} \tag{20}$$

where $\mathcal{H}_{\Gamma_d}$ is the Heaviside step function centred at the discontinuity. The decomposition in eq. (20) has a structure similar to the interpolation in eq. (19),

e.g. [17]. Accordingly, the partition-of-unity property of finite element shape functions enables the direct incorporation of discontinuities, including cracks and shear bands, in finite element models such that the discontinuous character of cracks and shear bands is preserved. With the standard small-strain assumption that the strain-rate field of the solid, $\boldsymbol{\epsilon}_s$, is derived from the symmetric part of the gradient of the velocity field, we obtain:

$$\dot{\boldsymbol{\epsilon}}_s = \nabla^s \bar{\mathbf{v}}_s + \mathcal{H}_{\Gamma_d} \nabla^s \tilde{\mathbf{v}}_s + \delta_{\Gamma_d} (\tilde{\mathbf{v}}_s \otimes \mathbf{n}_{\Gamma_d})^s \tag{21}$$

with the superscript s denoting the symmetric part of the gradient operator.

With respect to the pore fluid, we consider the case that a diaphragm with a permeability k_d is placed at the discontinuity in the displacement. As a consequence, the fluid pressure can be discontinuous across Γ_d and, similar to eq. (20), we have:

$$p = \bar{p} + \mathcal{H}_{\Gamma_d} \tilde{p} \tag{22}$$

It is noted that this assumption is different from that of Armero and Callari [18], who adopt a smooth pressure field (and therefore $p = \bar{p}$) and is also different from that of Larsson and Larsson [19], who assume that a regularised Dirac distribution is added to the continuous pressure field at the location of the discontinuity in the displacement field. For the fluid flow, gradients of the pressure need to be computed. Differentiating eq. (22), we obtain:

$$\nabla p = \nabla \bar{p} + \mathcal{H}_{\Gamma_d} \nabla \tilde{p} + \delta_{\Gamma_d} \tilde{p}\, \mathbf{n}_{\Gamma_d} \tag{23}$$

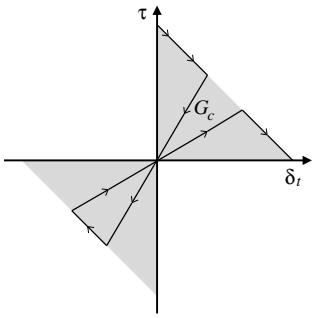


Fig. 2. Relation between relative sliding at the discontinuity and shear tractions

4 Constitutive equations

4.1 Models for the bulk

The effective stress increment in the solid skeleton, $\mathrm{d}\boldsymbol{\sigma}'_s$ is related to the strain increment $\mathrm{d}\boldsymbol{\epsilon}_s$ by an incrementally linear stress-strain relation for the solid skeleton,

$$\mathrm{d}\boldsymbol{\sigma}'_s = \bar{\mathbf{D}}^{tan} : \mathrm{d}\boldsymbol{\epsilon}_s \tag{24}$$

where $\bar{\mathbf{D}}^{tan}$ is the fourth–order tangent stiffness tensor of the solid material and the d - symbol denotes a small increment. Since the effective stress in the solid skeleton is related to the partial stress by $\boldsymbol{\sigma}'_s = \boldsymbol{\sigma}_s / n_s$, the above relation can be replaced by

$$\mathrm{d}\boldsymbol{\sigma}_s = \mathbf{D}^{tan} : \mathrm{d}\boldsymbol{\epsilon}_s \tag{25}$$

where the notation $\mathbf{D}^{tan} = n_s \bar{\mathbf{D}}^{tan}$ has been used. In the examples, a linear-elastic behaviour of the bulk material has been assumed, and we have set $\mathbf{D}^{tan} = \mathbf{D}$, the linear-elastic stiffness tensor.

For the flow of the pore fluid, Darcy's relation for isotropic media is assumed to hold,

$$n_f(\mathbf{v}_f - \mathbf{v}_s) = -k_f \nabla p \tag{26}$$

with k_f the permeability coefficient of the porous medium. For loading situations in which high strain rates play a significant role, Darcy's relation can be extended with a so-called dynamic seepage term [16, 20], which results in:

$$n_f(\mathbf{v}_f - \mathbf{v}_s) = -k_f \left(\nabla p + \rho_f \frac{\partial \mathbf{v}_f}{\partial t} \right) \tag{27}$$

In line with the earlier assumption to neglect the gravity acceleration, this term has also been omitted here. In practical situations, following the assumption $\frac{\partial \mathbf{v}_s}{\partial t} \approx \frac{\partial \mathbf{v}_f}{\partial t}$ for relatively slow dynamic loadings, eq. (27) is often approximated by

$$n_f(\mathbf{v}_f - \mathbf{v}_s) = -k_f \left(\nabla p + \rho_f \frac{\partial \mathbf{v}_s}{\partial t} \right) \tag{28}$$

4.2 Interface behaviour

At the discontinuity Γ_d a discrete relation holds between the interface tractions $\mathbf{t}_d$ and the relative displacements $\boldsymbol{\delta}$:

$$\mathbf{t}_d = \mathbf{t}_d(\boldsymbol{\delta}, \kappa) \tag{29}$$

with κ a history parameter. After linearisation, necessary to use a tangential stiffness matrix in an incremental-iterative solution procedure, one obtains:

$$\dot{\mathbf{t}}_d = \mathbf{T}\dot{\boldsymbol{\delta}} \tag{30}$$

with $\mathbf{T}$ the material tangent stiffness matrix of the discrete traction-separation law:

$$\mathbf{T} = \frac{\partial \mathbf{t}_d}{\partial \boldsymbol{\delta}} + \frac{\partial \mathbf{t}_d}{\partial \kappa} \frac{\partial \kappa}{\partial \boldsymbol{\delta}} \tag{31}$$

A first possibility that has been used in the example calculations for shear band initiation is the use of a maximum shear stress criterion in the spirit

of Tresca. With the resolved shear stress τ, a shear-band like discontinuity is created when the criterion equals the critical value τ_c: $\tau = \tau_c$. The orientation of the interface is such that it maximises the shear stress. In this orientation, the shear stress $\tau = \frac{|\sigma_1 - \sigma_2|}{2}$, σ_1, σ_2 being the principal stresses. A maximum shear stress nucleation criterion is primarily applicable when compressive stress states around the discontinuity prevail, such as in rocks and soils. Then, the failure mode will only involve sliding at the discontinuity, but no crack opening. For this reason, in the example calculations only degrees-of-freedom that describe this sliding mode have been added to the finite element model, which is different from earlier shear-band simulations (for single-phase media) [5, 6, 12] that have exploited the partition-of-unity property of finite element shape functions, but is similar to [9]. Dilatancy in the shear band can be incorporated when, in addition to the tangential degrees-of-freedom, during propagation extra degrees-of-freedom are activated which are normal to the shear band.

A key element is the presence of a mode-II fracture energy, $\mathcal{G}_c^{II}$, which governs the shear band evolution and enters the interface constitutive relation (29) in addition to the shear strength τ_c. It is defined as the work needed to create a unit area of fully developed shear band, e.g. [21]:

$$\mathcal{G}_c^{II} = \int_{\delta_t=0}^{\infty} \tau \mathrm{d}\delta_t \tag{32}$$

with τ the shear stress across the shear band, and δ_t the relative sliding between both faces of the shear band. $\mathcal{G}_c^{II}$ equals the area under the decohesion curves shown in Figure 2.

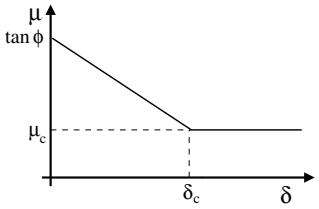


Fig. 3. Relation between relative sliding at the discontinuity and friction coefficient

Alternatively, a Coulomb criterion for local inception of the shear band has been used in the examples at the end of this paper. In this criterion nucleation starts when

$$\tau = \tau_{coh} + \tau_{fr} \tag{33}$$

with

$$\tau_{coh} = c_0$$

the cohesive contribution and c_0 the cohesion. Decohesion is governed by the fracture energy $\mathcal{G}_c^{II}$, similar to the Tresca-like criterion, cf. eq. (32). τ_{fr} is the

frictional contribution, which is defined as a function of the traction normal to the discontinuity and the effective friction coefficient μ:

$$\tau_{fr} = \mu \mathbf{n}_{\Gamma_d} \cdot \boldsymbol{\sigma} \cdot \mathbf{n}_{\Gamma_d}$$

The effective friction coefficient has a virgin value $\mu = \tan\phi$, with ϕ the friction angle. The vector $\mathbf{n}_{\Gamma_d}$ is such that it is normal to the critical plane where Coulomb's criterion for incipient shear failure is satisfied. A frictional softening relation models the microstructure evolution of the solid grains in the interface. The particular relation used in the example calculations is shown in Figure 3. In it, μ_c is the threshold value and δ_c the relative sliding at this value. It is assumed that $\delta_c = 2\mathcal{G}_c^{II}/c_0$, which equals the value defined by the cohesive softening relation.

As with the Tresca-like criterion, it is assumed that the failure mode only involves sliding. Possible dilatancy effects are not included in the kinematics of the discontinuity. For this reason, one can also now suffice by adding only degrees-of-freedom to the finite element model that describe the discrete sliding mode. It is interesting to note that, unlike in non-associated plasticity, the resulting stiffness matrix remains symmetric.

A discrete equivalent of Darcy's relation is now defined for the fluid flow $\mathbf{q}_d$ at the discontinuity as:

$$\mathbf{n}_{\Gamma_d} \cdot \mathbf{q}_d = -k_d(p^+ - p^-) = -k_d\, \tilde{p}\,|_{\mathbf{x}\in\Gamma_d} \tag{34}$$

where k_d is the permeability of the diaphragm that has been assumed to coincide with the displacement discontinuity Γ_d and p^+ and p^- are the pressures in the Ω^+ and Ω^- domains, respectively. For an impervious boundary, $k_d = 0$, which implies that $\mathbf{n}_{\Gamma_d} \cdot \mathbf{q}_d = 0$ according to eq. (34). Conversely, ideal permeability requires that $k_d \to \infty$, so that $\mathbf{n}_{\Gamma_d} \cdot \mathbf{q}_d$ can only be bounded if $p^+ - p^- = 0$, which implies that no discontinuity can exist in the pressure field and the formulation of Armero and Callari [18] is retrieved.

5 Numerical elaboration

5.1 Weak forms

To arrive at the weak form of the balance equations, we multiply the momentum balance (4) and the mass balance (15) by test functions for the velocities of the skeleton and for the pressures. In the spirit of a standard Bubnov-Galerkin approach, they are assumed to be of the following format:

$$\boldsymbol{\eta} = \bar{\boldsymbol{\eta}} + \mathcal{H}_{\Gamma_d}\tilde{\boldsymbol{\eta}} \tag{35}$$

for the velocities, and

$$\zeta = \bar{\zeta} + \mathcal{H}_{\Gamma_d}\tilde{\zeta} \tag{36}$$

for the pressures. Substitution into eqs (4) and (15), and integrating over the domain Ω leads to the corresponding weak forms:

$$\int_{\Omega} (\bar{\boldsymbol{\eta}} + \mathcal{H}_{\Gamma_d}\tilde{\boldsymbol{\eta}}) \cdot \left(\nabla \cdot \boldsymbol{\sigma} - \rho \frac{\partial \mathbf{v}_s}{\partial t} \right) \mathrm{d}\Omega = 0 \tag{37}$$

and

$$\int_{\Omega} (\bar{\zeta} + \mathcal{H}_{\Gamma_d}\tilde{\zeta}) \left(\alpha \nabla \cdot \mathbf{v}_s + n_f \nabla \cdot (\mathbf{v}_f - \mathbf{v}_s) + \frac{1}{Q} \frac{\partial p}{\partial t} \right) \mathrm{d}\Omega = 0 \tag{38}$$

Using the standard procedure of applying the divergence theorem, using the external boundary conditions (16) and (17), eliminating the Heaviside functions by changing the integration domain from Ω to Ω^+, eliminating the Dirac delta functions by transforming the volume integral into a surface integral, and introducing the shorter notation of a superimposed dot for $\partial/\partial t$, the balance equations take the form:

$$\begin{aligned} \int_{\Omega} \rho \bar{\boldsymbol{\eta}} \cdot \dot{\mathbf{v}}_s \mathrm{d}\Omega + \int_{\Omega^+} \rho \tilde{\boldsymbol{\eta}} \cdot \dot{\mathbf{v}}_s \mathrm{d}\Omega + \\ \int_{\Omega} (\nabla \cdot \bar{\boldsymbol{\eta}}) \cdot \boldsymbol{\sigma} \mathrm{d}\Omega + \int_{\Omega^+} (\nabla \cdot \tilde{\boldsymbol{\eta}}) \cdot \boldsymbol{\sigma} \mathrm{d}\Omega + \\ \int_{\Gamma_d} \tilde{\boldsymbol{\eta}} \cdot \mathbf{t}_d \mathrm{d}\Omega = \int_{\Gamma} (\bar{\boldsymbol{\eta}} + \mathcal{H}_{\Gamma_d}\tilde{\boldsymbol{\eta}}) \cdot \mathbf{t}_p \, \mathrm{d}\Gamma \end{aligned} \tag{39}$$

and

$$\begin{aligned} - \int_{\Omega} k_f \rho_f \nabla \bar{\zeta} \cdot \dot{\mathbf{v}}_s \mathrm{d}\Omega - \int_{\Omega^+} k_f \rho_f \nabla \tilde{\zeta} \cdot \dot{\mathbf{v}}_s \mathrm{d}\Omega \\ - \int_{\Omega} \alpha \bar{\zeta} \nabla \cdot \mathbf{v}_s \mathrm{d}\Omega - \int_{\Omega^+} \alpha \tilde{\zeta} \nabla \cdot \mathbf{v}_s \mathrm{d}\Omega \\ - \int_{\Omega} k_f \nabla \bar{\zeta} \cdot \nabla p \mathrm{d}\Omega - \int_{\Omega^+} k_f \nabla \tilde{\zeta} \cdot \nabla p \mathrm{d}\Omega - \int_{\Gamma_d} \tilde{\zeta} \mathbf{n}_{\Gamma_d} \cdot \mathbf{q}_d \, \mathrm{d}\Gamma \\ - \int_{\Omega} \bar{\zeta} Q^{-1} \dot{p} \, \mathrm{d}\Omega - \int_{\Omega^+} \tilde{\zeta} Q^{-1} \dot{p} \, \mathrm{d}\Omega = \int_{\Gamma} (\bar{\zeta} + \mathcal{H}_{\Gamma_d}\tilde{\zeta}) q_p \, \mathrm{d}\Gamma \end{aligned} \tag{40}$$

where for the derivation of the latter equation also Darcy's relation (28) has been employed.

5.2 Discretisations

We now switch to matrix-vector notation and discretise the trial functions $\mathbf{v}_s$ and p and the test functions $\boldsymbol{\eta}$ and ζ as:

$$\begin{aligned} \mathbf{v}_s &= \mathbf{N}(\dot{\bar{\mathbf{a}}} + \mathcal{H}_{\Gamma_d}\dot{\tilde{\mathbf{a}}}) \\ p &= \mathbf{H}(\bar{\mathbf{p}} + \mathcal{H}_{\Gamma_d}\tilde{\mathbf{p}}) \\ \boldsymbol{\eta} &= \mathbf{N}(\bar{\mathbf{w}} + \mathcal{H}_{\Gamma_d}\tilde{\mathbf{w}}) \\ \zeta &= \mathbf{H}(\bar{\mathbf{z}} + \mathcal{H}_{\Gamma_d}\tilde{\mathbf{z}}) \end{aligned} \tag{41}$$

Eqs (41) can be inserted into eqs (39) and (40) to obtain the semi-discrete form under the requirement that the result holds for all admissible $\bar{\mathbf{w}}$, $\bar{\mathbf{z}}$, $\tilde{\mathbf{w}}$ and $\tilde{\mathbf{z}}$.

The semi-discrete initial value problem is second order in time with respect to the displacement variables and first order for the fluid pore pressure. Yet, the same integration scheme will be used for both variables, in particular the Newmark method commonly used in structural dynamics. Let $\mathbf{y}$ denote an array which is a function of time. At the discrete time instant t_n its value is $\mathbf{y}_n$. Under the assumption of time continuity, the updating equations of the Newmark method are:

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \Delta t\dot{\mathbf{y}}_n + (\frac{1}{2} - \beta)\Delta t^2\ddot{\mathbf{y}}_n + \beta\Delta t^2\ddot{\mathbf{y}}_{n+1} \tag{42}$$

$$\dot{\mathbf{y}}_{n+1} = \dot{\mathbf{y}}_n + (1-\gamma)\Delta t\ddot{\mathbf{y}}_n + \gamma\Delta t\ddot{\mathbf{y}}_{n+1} \tag{43}$$

with β, γ the parameters of the time integration scheme. The equations can be recast as:

$$\begin{aligned} \ddot{\mathbf{y}}_{n+1} &= \alpha_0(\mathbf{y}_{n+1} - \mathbf{y}_n) - \alpha_2\dot{\mathbf{y}}_n - \alpha_4\ddot{\mathbf{y}}_n \\ \dot{\mathbf{y}}_{n+1} &= \alpha_1(\mathbf{y}_{n+1} - \mathbf{y}_n) - \alpha_3\dot{\mathbf{y}}_n - \alpha_5\ddot{\mathbf{y}}_n \end{aligned} \tag{44}$$

with

$$\alpha_0 = \frac{1}{\beta\Delta t^2} \;\;,\;\; \alpha_2 = \frac{1}{\beta\Delta t} \;\;,\;\; \alpha_4 = \frac{1}{2\beta} - 1$$
$$\alpha_1 = \frac{\gamma}{\beta\Delta t} \;\;,\;\; \alpha_3 = \frac{\gamma}{\beta} - 1 \;\;,\;\; \alpha_5 = (\frac{\gamma}{2\beta} - 1)\Delta t$$

For future use we also list the expressions for the variations that can be derived from expressions (44):

$$\delta\ddot{\mathbf{y}}_{n+1} = \alpha_0\delta\mathbf{y}_{n+1} \;\;,\;\; \delta\dot{\mathbf{y}}_{n+1} = \alpha_1\delta\mathbf{y}_{n+1} \tag{45}$$

Application of the time integration scheme (44) to semi-discrete balance equations results in a set of coupled, discrete equations, which is nonlinear. Therefore, an iterative solution procedure has to be applied within each time step Δt. When using the Newton-Raphson method, as has been done in the ensuing examples, and exploiting the variations defined in eqs (45), one obtains a sequence of linearised problems, which for implementation purposes are conveniently be cast in a matrix-vector format:

$$\begin{bmatrix} \alpha_0\mathbf{M}_{\bar{a}\bar{a}} + \mathbf{K}_{\bar{a}\bar{a}} & \alpha_0\mathbf{M}_{\bar{a}\tilde{a}} + \mathbf{K}_{\bar{a}\tilde{a}} & \mathbf{K}_{\bar{a}\bar{p}} & \mathbf{K}_{\bar{a}\tilde{p}} \\ \alpha_0\mathbf{M}_{\tilde{a}\bar{a}} + \mathbf{K}_{\tilde{a}\bar{a}} & \alpha_0\mathbf{M}_{\tilde{a}\tilde{a}} + \mathbf{K}_{\tilde{a}\tilde{a}} & \mathbf{K}_{\tilde{a}\bar{p}} & \mathbf{K}_{\tilde{a}\tilde{p}} \\ \alpha_0\mathbf{M}_{\bar{p}\bar{a}} + \alpha_1\mathbf{K}^{\mathrm{T}}_{\bar{a}\bar{p}} & \alpha_0\mathbf{M}_{\bar{p}\tilde{a}} + \alpha_1\mathbf{K}^{\mathrm{T}}_{\tilde{a}\bar{p}} & \alpha_1\mathbf{M}_{\bar{p}\bar{p}} + \mathbf{K}_{\bar{p}\bar{p}} & \alpha_1\mathbf{M}_{\bar{p}\tilde{p}} + \mathbf{K}_{\bar{p}\tilde{p}} \\ \alpha_0\mathbf{M}_{\tilde{p}\bar{a}} + \alpha_1\mathbf{K}^{\mathrm{T}}_{\bar{a}\tilde{p}} & \alpha_0\mathbf{M}_{\tilde{p}\tilde{a}} + \alpha_1\mathbf{K}^{\mathrm{T}}_{\tilde{a}\tilde{p}} & \alpha_1\mathbf{M}_{\tilde{p}\bar{p}} + \mathbf{K}_{\tilde{p}\bar{p}} & \alpha_1\mathbf{M}_{\tilde{p}\tilde{p}} + \mathbf{K}_{\tilde{p}\tilde{p}} \end{bmatrix} \begin{pmatrix} \mathrm{d}\bar{\mathbf{a}} \\ \mathrm{d}\tilde{\mathbf{a}} \\ \mathrm{d}\bar{\mathbf{p}} \\ \mathrm{d}\tilde{\mathbf{p}} \end{pmatrix} = \begin{pmatrix} \mathbf{f}^*_{\bar{a}} \\ \mathbf{f}^*_{\tilde{a}} \\ \mathbf{f}^*_{\bar{p}} \\ \mathbf{f}^*_{\tilde{p}} \end{pmatrix} \tag{46}$$

with the external force vectors:

$$\mathbf{f}_{\bar{a}}^{ext} = \int_{\Gamma} \mathbf{N}^{\mathrm{T}} \mathbf{t}_p \mathrm{d}\Gamma \ \ , \ \ \mathbf{f}_{\tilde{a}}^{ext} = \int_{\Gamma} \mathcal{H}_{\Gamma_d} \mathbf{N}^{\mathrm{T}} \mathbf{t}_p \mathrm{d}\Gamma$$
$$\mathbf{f}_{\bar{p}}^{ext} = \int_{\Gamma} \mathbf{H}^{\mathrm{T}} q_p \mathrm{d}\Gamma \ \ , \ \ \mathbf{f}_{\tilde{p}}^{ext} = \int_{\Gamma} \mathcal{H}_{\Gamma_d} \mathbf{H}^{\mathrm{T}} q_p \mathrm{d}\Gamma$$

the internal force vectors:

$$\mathbf{f}_{\bar{a}}^{int} = \int_{\Omega} \mathbf{B}^{\mathrm{T}} \boldsymbol{\sigma} \mathrm{d}\Omega$$
$$\mathbf{f}_{\tilde{a}}^{int} = \int_{\Omega^+} \mathbf{B}^{\mathrm{T}} \boldsymbol{\sigma} \mathrm{d}\Omega + \int_{\Gamma_d} \mathbf{N}^{\mathrm{T}} \mathbf{t}_d \mathrm{d}\Gamma$$

with $\mathbf{B} = \nabla \mathbf{N}$ and, for two dimensions, $\mathbf{m} = [1, 1, 0]$. The mass matrices:

$$\mathbf{M}_{\bar{a}\bar{a}} = \int_{\Omega} \rho \mathbf{N}^{\mathrm{T}} \mathbf{N} \mathrm{d}\Omega \ \ , \ \mathbf{M}_{\tilde{a}\bar{a}} = \mathbf{M}_{\bar{a}\tilde{a}} = \mathbf{M}_{\tilde{a}\tilde{a}} = \int_{\Omega^+} \rho \mathbf{N}^{\mathrm{T}} \mathbf{N} \mathrm{d}\Omega$$
$$\mathbf{M}_{\bar{p}\bar{a}} = -\int_{\Omega} k_f \rho_f \nabla \mathbf{H}^{\mathrm{T}} \mathbf{N} \mathrm{d}\Omega \ \ , \ \mathbf{M}_{\tilde{p}\bar{a}} = \mathbf{M}_{\bar{p}\tilde{a}} = \mathbf{M}_{\tilde{p}\tilde{a}} = -\int_{\Omega^+} k_f \rho_f \nabla \mathbf{H}^{\mathrm{T}} \mathbf{N} \mathrm{d}\Omega$$
$$\mathbf{M}_{\bar{p}\bar{p}} = -\int_{\Omega} Q^{-1} \mathbf{H}^{\mathrm{T}} \mathbf{H} \mathrm{d}\Omega \ \ , \ \mathbf{M}_{\tilde{p}\bar{p}} = \mathbf{M}_{\bar{p}\tilde{p}} = \mathbf{M}_{\tilde{p}\tilde{p}} = -\int_{\Omega^+} Q^{-1} \mathbf{H}^{\mathrm{T}} \mathbf{H} \mathrm{d}\Omega$$

the stiffness matrices:

$$\mathbf{K}_{\bar{a}\bar{p}} = -\int_{\Omega} \alpha \mathbf{B}^{\mathrm{T}} \mathbf{m} \mathbf{H} \mathrm{d}\Omega \ \ , \ \ \mathbf{K}_{\tilde{a}\bar{p}} = \mathbf{K}_{\bar{a}\tilde{p}} = \mathbf{K}_{\tilde{a}\tilde{p}} = -\int_{\Omega^+} \alpha \mathbf{B}^{\mathrm{T}} \mathbf{m} \mathbf{H} \mathrm{d}\Omega$$
$$\mathbf{K}_{\bar{p}\bar{p}} = -\int_{\Omega} k_f \nabla \mathbf{H}^{\mathrm{T}} \nabla \mathbf{H} \mathrm{d}\Omega \ \ , \ \ \mathbf{K}_{\tilde{p}\bar{p}} = \mathbf{K}_{\bar{p}\tilde{p}} = -\int_{\Omega^+} k_f \nabla \mathbf{H}^{\mathrm{T}} \nabla \mathbf{H} \mathrm{d}\Omega$$
$$\mathbf{K}_{\tilde{p}\tilde{p}} = -\int_{\Omega^+} k_f \nabla \mathbf{H}^{\mathrm{T}} \nabla \mathbf{H} \mathrm{d}\Omega - \int_{\Gamma_d} k_d \mathbf{H}^{\mathrm{T}} \mathbf{H} \mathrm{d}\Gamma$$
$$\mathbf{K}_{\bar{a}\bar{a}} = \int_{\Omega} \mathbf{B}^{\mathrm{T}} \mathbf{D} \mathbf{B} \mathrm{d}\Omega \ \ , \ \ \mathbf{K}_{\tilde{a}\bar{a}} = \mathbf{K}_{\bar{a}\tilde{a}} = \int_{\Omega^+} \mathbf{B}^{\mathrm{T}} \mathbf{D} \mathbf{B} \mathrm{d}\Omega$$
$$\mathbf{K}_{\tilde{a}\tilde{a}} = \int_{\Omega^+} \mathbf{B}^{\mathrm{T}} \mathbf{D} \mathbf{B} \mathrm{d}\Omega + \int_{\Gamma_d} \mathbf{N}^{\mathrm{T}} \mathbf{T} \mathbf{N} \mathrm{d}\Gamma$$

and the arrays at the right-hand side:

$$\mathbf{f}_{\bar{a}}^{*} = \mathbf{f}_{\bar{a}}^{ext} - (\mathbf{f}_{\bar{a}}^{int})^i - \alpha_0 \mathbf{M}_{\bar{a}\bar{a}} \ddot{\bar{\mathbf{a}}}_{n+1}^i - \alpha_0 \mathbf{M}_{\bar{a}\tilde{a}} \ddot{\tilde{\mathbf{a}}}_{n+1}^i$$
$$\mathbf{f}_{\tilde{a}}^{*} = \mathbf{f}_{\tilde{a}}^{ext} - (\mathbf{f}_{\tilde{a}}^{int})^i - \alpha_0 \mathbf{M}_{\tilde{a}\bar{a}} \ddot{\bar{\mathbf{a}}}_{n+1}^i - \alpha_0 \mathbf{M}_{\tilde{a}\tilde{a}} \ddot{\tilde{\mathbf{a}}}_{n+1}^i$$
$$\mathbf{f}_{\bar{p}}^{*} = \mathbf{f}_{\bar{p}}^{ext} - \alpha_0 \mathbf{M}_{\bar{p}\bar{a}} \ddot{\bar{\mathbf{a}}}_{n+1}^i - \alpha_0 \mathbf{M}_{\bar{p}\tilde{a}} \ddot{\tilde{\mathbf{a}}}_{n+1}^i - \alpha_1 \mathbf{M}_{\bar{p}\bar{p}} \dot{\bar{\mathbf{p}}}_{n+1}^i - \alpha_1 \mathbf{M}_{\bar{p}\tilde{p}} \dot{\tilde{\mathbf{p}}}_{n+1}^i$$
$$-\alpha_1 \mathbf{K}_{\bar{a}\bar{p}}^{\mathrm{T}} \dot{\bar{\mathbf{a}}}_{n+1}^i - \alpha_1 \mathbf{K}_{\tilde{a}\bar{p}}^{\mathrm{T}} \dot{\tilde{\mathbf{a}}}_{n+1}^i$$
$$\mathbf{f}_{\tilde{p}}^{*} = \mathbf{f}_{\tilde{p}}^{ext} + -\alpha_0 \mathbf{M}_{\tilde{p}\bar{a}} \ddot{\bar{\mathbf{a}}}_{n+1}^i - \alpha_0 \mathbf{M}_{\tilde{p}\tilde{a}} \dot{\tilde{\mathbf{a}}}_{n+1}^i - \alpha_1 \mathbf{M}_{\tilde{p}\bar{p}} \dot{\bar{\mathbf{p}}}_{n+1}^i - \alpha_1 \mathbf{M}_{\tilde{p}\tilde{p}} \dot{\tilde{\mathbf{p}}}_{n+1}^i$$
$$-\alpha_1 \mathbf{K}_{\bar{a}\tilde{p}}^{\mathrm{T}} \dot{\bar{\mathbf{a}}}_{n+1}^i - \alpha_1 \mathbf{K}_{\tilde{a}\tilde{p}}^{\mathrm{T}} \dot{\tilde{\mathbf{a}}}_{n+1}^i$$

where the superscript signifies that the corresponding quantity has to be evaluated at iteration i. The quantities $\ddot{\bar{\mathbf{a}}}_{n+1}^i, \ddot{\tilde{\mathbf{a}}}_{n+1}^i, \dot{\bar{\mathbf{a}}}_{n+1}^i, \dot{\tilde{\mathbf{a}}}_{n+1}^i, \dot{\bar{\mathbf{p}}}_{n+1}^i, \dot{\tilde{\mathbf{p}}}_{n+1}^i$ are evaluated using eqs (44).

The stiffness matrix of eq. (46) is not symmetric. Symmetry can be restored by multiplying the third and the fourth row of submatrices by α_1^{-1} and omitting the contributions in the tangent stiffness matrix that are due to the dynamic seepage term – the submatrices $\mathbf{M}_{\bar{p}\bar{a}}$, $\mathbf{M}_{\bar{p}\tilde{a}}$, $\mathbf{M}_{\tilde{p}\bar{a}}$ and $\mathbf{M}_{\tilde{p}\tilde{a}}$. Since the corresponding terms are retained in the right-hand side, the results are not affected, only the convergence speed of the iterative procedure.

5.3 Stress computation at the tip

The nucleation criterion requires the determination of the stresses at the tip of the discontinuity. Unfortunately, the stresses vary strongly in the vicinity of the tip and an accurate estimate of them is difficult to obtain. In the present case, the stress in the bulk of the specimen is almost homogeneous except for a small area around the tip, which exacerbates the problem. Following Wells [5] and Jirasek [23] we use a smoothing of the stresses around the tip and compute the stress at the tip by the following nonlocal-like procedure:

$$\boldsymbol{\sigma}_{tip} = \frac{\int_\Omega w\boldsymbol{\sigma}\ \mathrm{d}\Omega}{\int_\Omega w\ \mathrm{d}\Omega} \tag{47}$$

where w is a Gaussian weight function:

$$w = e^{-r^2/2l^2}$$

with r the distance to the tip, and l a characteristic length which defines the size of region of influence of the stress. Because of the nearly homogeneous stress state in the specimen, a small value of l is desired, preferably in the same order of magnitude as the characteristic element length. This is accomplished in the following manner. By virtue of the linear behaviour of the solid phase in the bulk, a separate, independent integration domain can be defined, which follows the tip during propagation. This domain contains integration cells smaller than those of the mesh used in the discretisation – typically their length is in the order of 15-20% of the element size. Moreover, a higher-order Gaussian quadrature is used over this domain, which results in a very accurate determination of the tip stress.

6 Example calculations

All results of the computations are based for the same two-dimensional specimen with a width $w = 0.04$ m and a height $H = 0.1$ m, see also Figure 4, which is loaded under plane-strain conditions. The sides are traction free and the external loading is applied via an imposed constant velocity $V_0 = -10^{-3}$ m/s. Undrained conditions have been imposed on the entire boundary of the specimen, because fast transient phenomena have been considered. The solid constituent is assumed to behave in a linear elastic manner

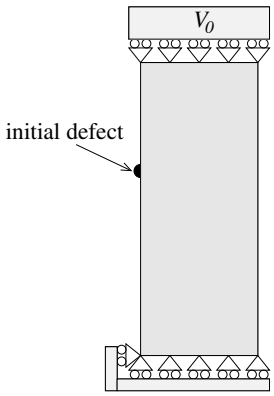


Fig. 4. Geometry and boundary conditions

with a Young's modulus $E = 20$ GPa and a Poisson's ratio $\nu = 0.35$. The absolute mass densities are $\rho'_s = \rho_s/n_s = 2000$ kg/m^3 for the solid phase and $\rho'_f = \rho_f/n_f = 1000$ kg/m^3 for the fluid phase, while the fluid fraction $n_f = 0.3$. The Biot coefficient α has been set equal to 1, the Biot modulus has been assigned a value $Q = 5.0$ GPa, while the bulk material was assumed to have a permeability $k_f = 10^{-14}$ m^3/Ns. The permeability of the diaphragm was assigned a value $k_d = 0.5 \cdot 10^{-14}$ m^2/Ns. Shear-band formation was triggered by a small imperfection, see Figure 4.

A structured mesh has been used and consists of 5841 four-noded elements with equal (bilinear) interpolations for the displacements and the pressure. The simulation is started using a time step of 0.4 s, which is small enough to accurately follow the pressure evolution and the near quasi-static behavior of the solid skeleton before the onset of the shear band. When the shear band starts to propagate, the phenomenon becomes dynamic and the time step size is reduced severely in order to properly capture the propagation of the stress wave. The parameters of the Newmark scheme are $\gamma = 0.5$ and $\beta = 0.25$.

6.1 Tresca-like initation criterion

The simulation for the Tresca-like nucleation criterion from which most of the results derive, has been obtained with the following parameters: the time step size during the shear-band formation equals 0.2 μs, nucleation traction $\tau_c = 50$ MPa and mode-II fracture energy $\mathcal{G}_c^{II} = 750$ J/m^2. The dynamic seepage term has not been taken into account in the analysis.

The evolution of the pore pressure field following the time t_0 at which the shear band starts to propagate, is shown in Figure 5. The scale has been chosen such that the white regions on the picture have a pore pressure below the cavitation pressure (here: -10^{+5}Pa). One observes that, initially, cavitation occurs only in the close vicinity of the discontinuity. When the shear band tip reaches the centre of the specimen, the level of the pore pressure above the

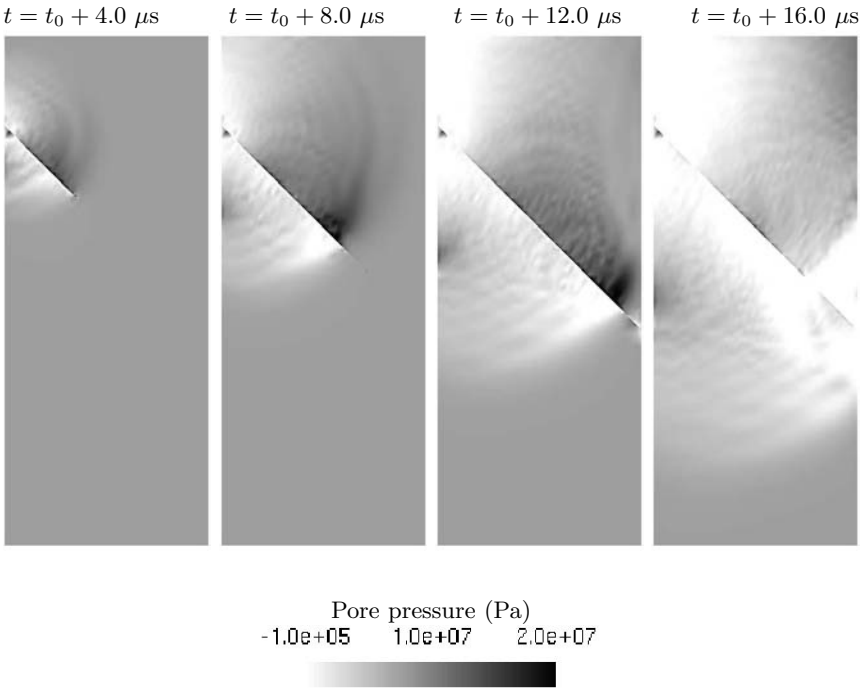


Fig. 5. Evolution of the pressure field for the Tresca criterion

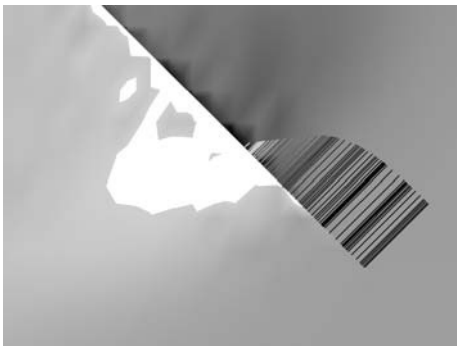


Fig. 6. Pressure field near the process zone and tractions at the discontinuity. The scale of the pressures is equal to that in Figure 5. The magnitudes of the tractions are proportional to the lengths of the bars

discontinuity increases and pore pressures that exceed the cavitation pressure develop over a larger region below the interface. Subsequently, this region follows the tip of the shear band, and when it reaches the right boundary of the specimen, the cavitation phenomena extend over the entire specimen. Because of the limitations of the model – a gas phase has not been modelled separately – the physical interpretation of the numerical results at this advanced stage of shear band propagation becomes questionable.

As illustrated by Figure 6, it seems that in the present simulations that utilise the Tresca-like criterion, cavitation is a consequence of local elastic unloading behind the process zone, where cohesive softening takes place. Indeed, Figure 6 shows the pressure field as well as the values of the shear tractions in the cohesive interface (with bars orthogonal to the discontinuity). The cavitation front appears to coincide with the transition zone between the damaged and the intact parts of the interface.

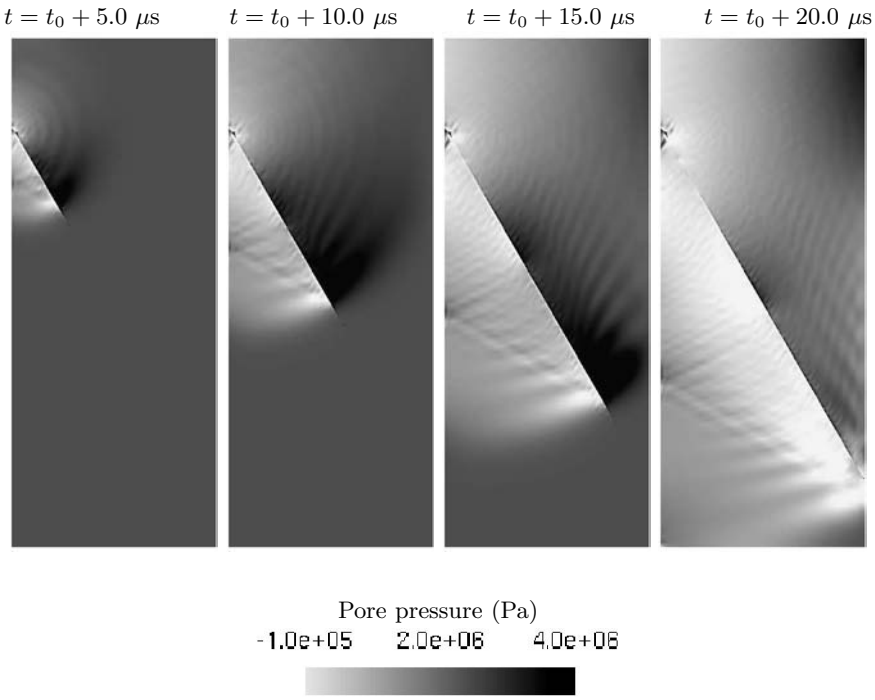


Fig. 7. Evolution of the pressure field for the Coulomb criterion with $\frac{\delta_c}{\Delta U_{ini}} = 0.024$

6.2 Coulomb initiation criterion

For the Coulomb criterion, the following model parameters have been used: cohesion $c_0 = 10.0$ MPa, friction angle $\phi = 30^o$. The threshold value in the

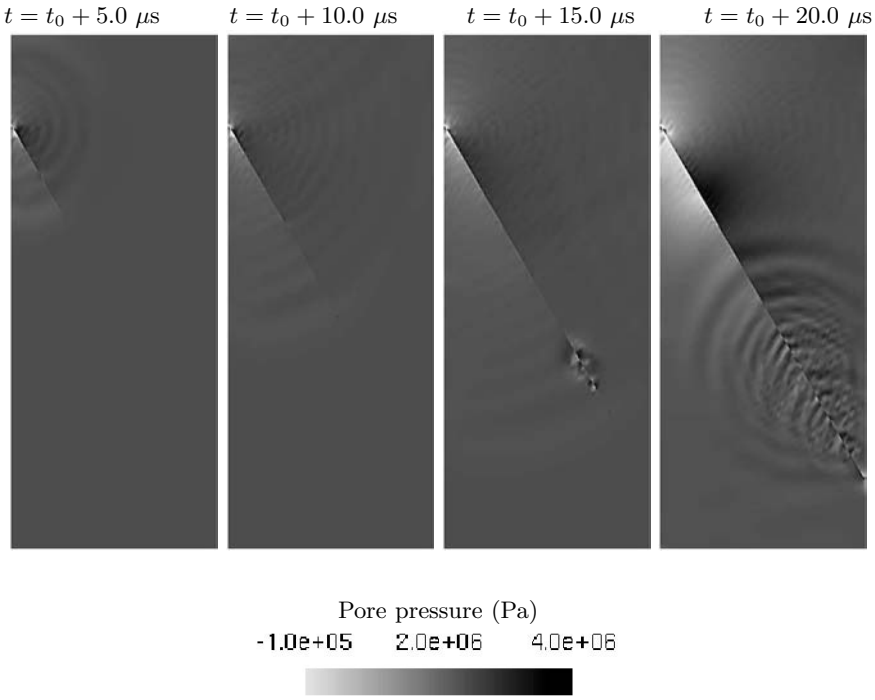


Fig. 8. Evolution of the pressure field for the Coulomb criterion with $\frac{\delta_c}{\Delta U_{ini}} = 0.16$

frictional softening law is $\mu_c = 0.2\ \tan\phi$. A parametric study has been carried out with respect to the influence of $\frac{\delta_c}{\Delta U_{ini}}$, where δ_c is the tangential displacement jump when the cohesive part of the traction has vanished, and ΔU_{ini} the value of the prescribed displacement at the top of the specimen at shear-band initiation. The simulations have been carried out using a time step size of 0.25 μs.

A first simulation has been carried obtained for a fracture energy $\mathcal{G}_c^{II} = 15\ \mathrm{J/m^2}$. In this case $\frac{\delta_c}{\Delta U_{ini}}$ equals 0.024. The evolution of the pressure field is shown in Figure 7. The results are quite similar to those obtained with the Tresca criterion since the values of $\frac{\delta_c}{\Delta U_{ini}}$ are close. Figure 10 shows that for this case the process zone, i.e. where the tractions do not vanish, is rather small. As a consequence, local unloadings are obtained behind this process zone and cavitation occurs only in a small zone behind the shear-band tip.

Increasing the value of $\frac{\delta_c}{\Delta U_{ini}}$ to 0.16, the length of the process zone becomes approximately equal to that of the shear band, see Figure 10. Now, local unloadings are not observed, but a global unloading occurs that can be associated with the strain localisation inside the shear band and the softening of the interface at the end of propagation. As a consequence, the pressure evolution on Figure 8 has no local cavitation zone. Indeed, cavitation is first obtained around the initiation locus and, subsequently, in the entire specimen.

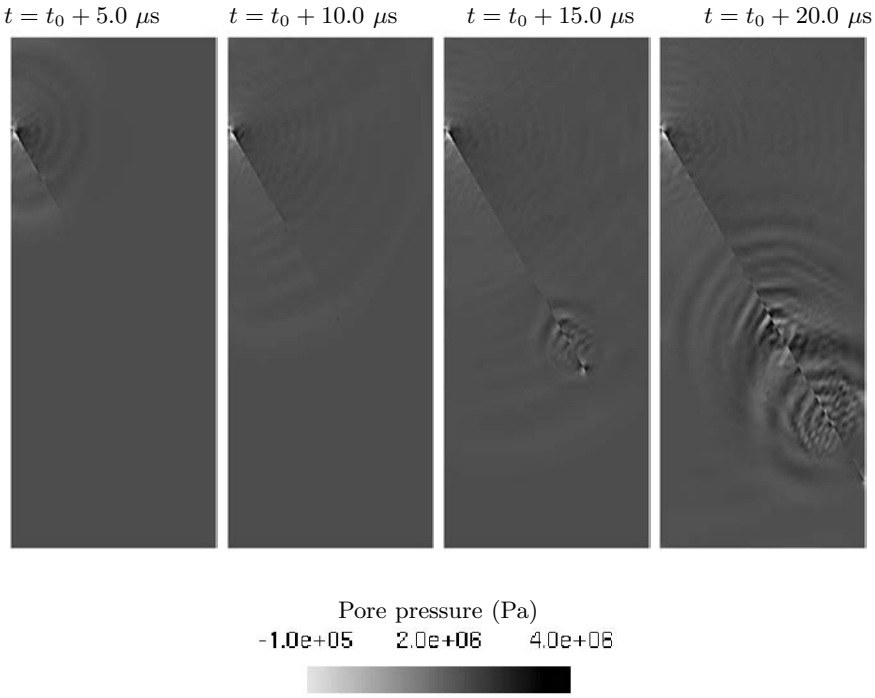


Fig. 9. Evolution of the pressure field for the Coulomb criterion with $\frac{\delta_c}{\Delta U_{ini}} = 0.24$

Instability patterns are observed on the last two figures that plot the pressure distribution.

Such instabilities, which are due to friction, are observed at a larger scale when $\frac{\delta_c}{\Delta U_{ini}}$ is increased further up to 0.24. Now, the shear band propagates because the stress is higher than the material strength, but the mechanical energy is not sufficient to damage the interface. Consequently, no localisation is obtained, but only frictional instabilities are observed. No cavitation is induced because of the absence of strain localisation.

7 Concluding remarks

In this contribution a numerical model has been elaborated which can capture discontinuities, e.g. cracks or shear bands, in a fluid-saturated medium. The representation of the discontinuity is truly discrete and unbiased by the discretisation. Moreover, the constitutive relations for the bulk and for the discontinuity can be specified independently, for the solid phase as well as for the fluid phase. Example calculations of dynamic shear band propagation have been presented with a Tresca-like and a Coulomb criterion for shear band initiation. The results show that the propagation of the shear band is strongly

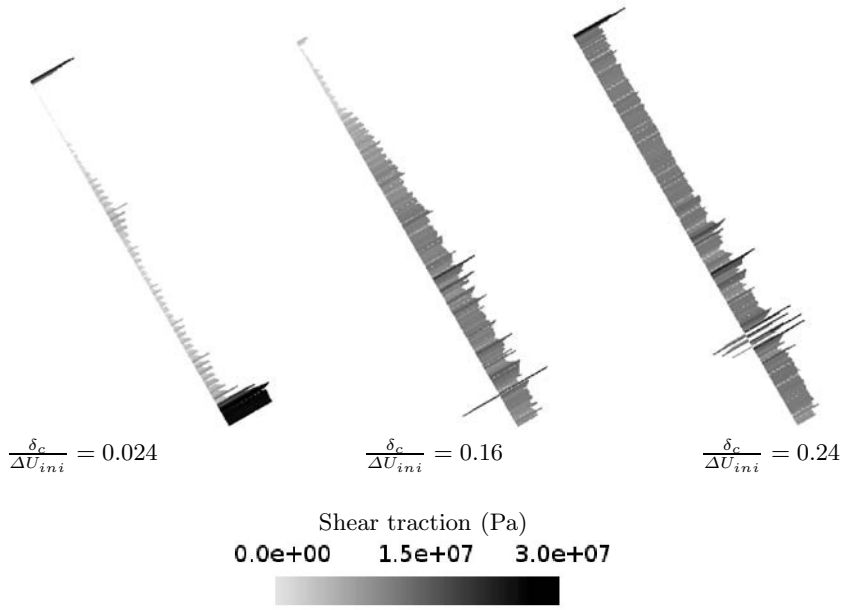


Fig. 10. Shear traction distribution at the end of the shear–band propagation for different values of $\frac{\delta_c}{\Delta U_{ini}}$

influenced by the constitutive assumptions in the discontinuity. Indeed, as highlighted by the results of the calculations, the cavitation phenomenon is triggered by unloading of the solid skeleton, which is a direct consequence of strain localisation and strongly depends on the constitutive model for the discontinuity.

8 Acknowledgements

This work was supported by the European Commission under contract HPRN-CT-2002-00198.

References

1. de Borst R (2004) Damage, material instabilities, and failure. In: Encyclopedia of Computational Mechanics. Wiley, Chichester
2. Babuska I, Melenk JM (1997) International Journal for Numerical Methods in Engineering 40:727–758
3. Belytschko T, Black T (1999) International Journal for Numerical Methods in Engineering 45:601–620

4. Moës N, Dolbow J, Belytschko T (1999) International Journal for Numerical Methods in Engineering 46:131–150
5. Wells GN, Sluys LJ (2001) International Journal for Numerical and Analytical Methods in Geomechanics 25:691–709
6. Wells GN, Sluys LJ, de Borst R (2002) International Journal for Numerical Methods in Engineering 53:1235–1256
7. Wells GN, de Borst R, Sluys LJ (2002) International Journal for Numerical Methods in Engineering 54:1333–1355
8. Remmers JJC, de Borst R, Needleman A (2003) Computational Mechanics 31:69–77
9. Samaniego E, Belytschko T (2005) International Journal for Numerical Methods in Engineering 62:1857–1872
10. Réthoré J, Gravouil A, Combescure A (2005) International Journal for Numerical Methods in Engineering 63:631–659
11. Réthoré J, Gravouil A, Combescure A (2005) International Journal for Numerical Methods in Engineering 64:260–284
12. Areias PMA, Belytschko T (2006) International Journal for Numerical Methods in Engineering 66:878–910
13. Terzaghi K (1943) Theoretical Soil Mechanics John Wiley & Sons, New York
14. Biot MA (1965) Mechanics of Incremental Deformations John Wiley & Sons, Chichester
15. Jouanna P, Abellan MA (1995) Generalized approach to heterogeneous media In: Modern Issues in Non-Saturated Soils. Springer-Verlag, Wien - New York, 1–128
16. Lewis RW, Schrefler BA (1998) The Finite Element Method in the Static and Dynamic Deformation and Consolidation of Porous Media, Second Edition John Wiley & Sons, Chichester
17. de Borst R, Réthoré J, Abellan MA (2006) Archive of Applied Mechanics 75:595–606
18. Armero F, Callari C (1999) International Journal for Numerical Methods in Engineering 46:1673–1698
19. Larsson J, Larsson R (2000) Mechanics of Cohesive-frictional Materials 5:565–582
20. Schrefler, B, Scotta R (2001) Computer Methods in Applied Mechanics and Engineering 190:3223–3246
21. Feenstra PH, de Borst R (1996) International Journal of Solids and Structures 33:707–730
22. Abellan MA, de Borst R (2006) Computer Methods in Applied Mechanics and Engineering 195:5011–5019
23. Jirasek M (1998) Embedded crack models for concrete fracture In: Computational Modelling of Concrete Structures. Balkema, Rotterdam, 291–300

Single Domain Quadrature Techniques for Discontinuous and Non-Linear Enrichments in Local Partion of Unity FEM

Giulio Ventura

Department of Structural Engineering and Geotechnics, Politecnico di Torino, I-10129 Torino, Italy; giulio.ventura@polito.it

Summary. The problem of the evaluation of the stiffness matrix for finite elements enriched by discontinuous/non-linear functions is investigated. If the introduction of discontinuities inside the elements through enrichment functions is nowadays well established by local partition of unity techniques, the evaluation of the element stiffness requires splitting the element into quadrature subcells where appropriate quadrature rules apply. To overcome this problem a technique is suggested, called polynomial mapping, based on replacing the enrichment function with polynomials having the same integral of the original function. These polynomials are function of the position of the discontinuity and are defined on the entire element domain, therefore avoiding the generation of quadrature subcells. The technique is applied to discontinuities in the displacement and strain and is introduced for regularized jumps in the displacement. An integration error analysis is shown in the latter case.

Key words: Extended finite element method, partition of unity, regularized displacement jump, quadrature, equivalent polynomial.

1 Introduction

Local Partition of Unity Finite Element Methods [1] (like the eXtended Finite Element Method) enrich the standard finite element by introducing suitable discontinuous and/or non-linear functions to best approximate the solution and to model arbitrary discontinuities independent of the mesh [2–5]. In this context a well-known problem is given by the numerical evaluation of the stiffness matrix in the elements intersected by the discontinuity surface and in the elements where the enrichment is highly non-linear. In fact, in this case, Gaussian quadrature is inadequate as large errors may arise. This is commonly solved by partitioning the elements to generate quadrature subcells [6, 7] where the integrands are continuous and differentiable, by high order Gauss rules or by adaptive quadrature with all the relevant computational costs. In this work the problem of the quadrature of discontinuous or

Alain Combescure et al. (eds.), IUTAM Symposium on Discretization Methods for Evolving Discontinuities, 343–361.

non-linear functions in the enriched elements is investigated. A technique is illustrated to eliminate the requirement of quadrature subcells generation. The technique extends the concepts introduced in [8] and it relies on the analytic study of the integrand of the element stiffness. Let the signed distance from the discontinuity surface be a level set function. Based on the nodal values of the level set, it is demonstrated that an equivalent polynomial function exists such that its integral gives the exact value of the original function integrated on subcells. The polynomial is defined in the whole element domain, so that it can be conveniently integrated by Gauss quadrature, and no quadrature subdomains need be defined. This has been developed, in the hypothesis of a linear discontinuity surface crossing completely an element, with reference to displacement jumps in one, two and three dimensions, as well as in the case of a discontinuity in the strain field due to a material interface. The case of a regularized Heaviside displacement jump is here analyzed as a first application to a continuous but non-polynomial and highly non-linear enrichment. The problem is formulated in an arbitrary number of dimensions and, for the sake of brevity, results are given in the one dimensional case: polynomials are given in closed form replacing the enrichment function and its derivative in the element stiffness evaluation. An error analysis for different quadrature strategies is performed and the results are reported and commented.

2 Analysis of the Problem

Consider a body described by a domain Ω with boundary $\partial\Omega$ and let $\mathbf{u}$ be the displacement field.

The boundary is partitioned into $\partial\Omega_u$ where the displacements are prescribed, $\mathbf{u} = \bar{\mathbf{u}}$, and $\partial\Omega_q$ where the traction $\mathbf{q}$ is given, such that $\partial\Omega_u \bigcap \partial\Omega_q = \varnothing$.

The local Partition of Unity approximation to the displacement field is:

$$\mathbf{u}(\mathbf{x}) = \sum_{I \in \mathcal{N}} N_I(\mathbf{x}) \left(\mathbf{u}_I + \mathbf{a}_I f_{\mathrm{e}}(\mathbf{x})\right) \tag{1}$$

where $\mathcal{N}$ are the nodes of the finite element mesh, $N_I(\mathbf{x})$ the finite element shape functions and the coefficients $\mathbf{u}_I$, $\mathbf{a}_I$ are to be determined through the application of a variational principle or weak form to the discrete displacement field (1).

Suppose the enrichment f_{e} is a function of the signed distance $d(\mathbf{x})$ from a given surface Γ. In the case of cracks Γ is the crack (discontinuity) surface.

The enrichment part is usually added only to the nodes of the elements intersected by the surface Γ or to a suitable neighbor of this surface, so that these elements are called *enriched elements*. The coefficients $\mathbf{a}_I$ are not included (i.e. set to zero) at nodes of unenriched elements.

As described in [3], the discontinuity in the displacement field given by a crack can be represented assuming as enrichment function f_{e} the generalized

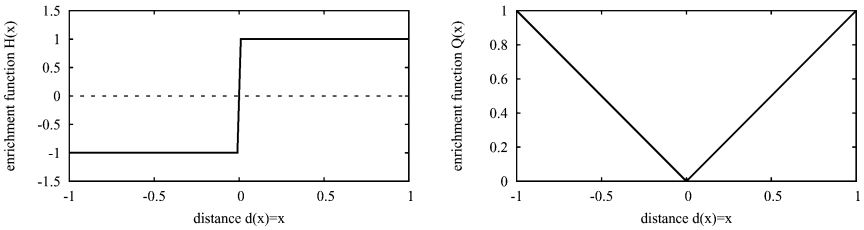


Fig. 1. Enrichment functions: (a) for crack problems, (b) for material discontinuity.

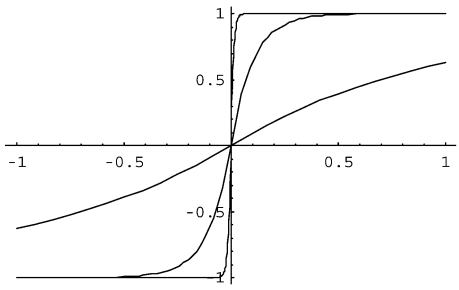


Fig. 2. Regularized Heaviside function plots for $\delta = 1.0,\ 0.1,\ 0.01$.

Heaviside step function, Fig. 1a

$$f_{\mathrm{e}}(\mathbf{x}) = H(\mathbf{x}) = \mathrm{sign}\,(d(\mathbf{x})) = \begin{cases} 1 \text{ if } d(\mathbf{x}) \geq 0 \\ -1 \text{ if } d(\mathbf{x}) < 0 \end{cases} \tag{2}$$

Material discontinuities, generating a discontinuity in the strain field, can be represented by the function, Fig. 1b

$$f_{\mathrm{e}}(\mathbf{x}) = Q(\mathbf{x}) = \mathrm{abs}\,(d(\mathbf{x})) = \begin{cases} d(\mathbf{x}) \text{ if } d(\mathbf{x}) \geq 0 \\ -d(\mathbf{x}) \text{ if } d(\mathbf{x}) < 0 \end{cases} \tag{3}$$

Finally, in recent works on shear bands and localization [9,10] regularized Heaviside functions have been considered, i.e. continuous and differentiable functions that tend to the Heaviside step function according to the value of a parameter. In the present work the following regularized Heaviside function will be considered, Fig. 2

$$f_{\mathrm{e}}(d(\mathbf{x})) = H_\rho(\mathbf{x}) = \int_0^{d(\mathbf{x})} e^{-\frac{|\xi|}{\rho}}\, d\xi = \mathrm{sign}(d(\mathbf{x}))(1 - e^{-\frac{|d(\mathbf{x})|}{\rho}})) \tag{4}$$

The traditional Gaussian quadrature used to evaluate the element stiffness matrices cannot be immediately applied when the above enrichment functions

are present, being it suitable only for polynomial functions. As mentioned in the Introduction, this is usually solved by splitting the element domain into two quadrature subdomains, determined by the discontinuity surface Γ, or by finer cell subdivisions and higher Gauss quadrature rules for non-polynomial and non-linear enrichments.

To study how the subdivision of the quadrature domain can be eliminated, it is useful to consider the symbolic form of the element stiffness matrix. In the following subsections this will be developed for the three listed enrichments (2), (3), (4), but the method may be extended to other enrichment functions.

The displacement and strain fields in a single enriched finite element can be written in the form

$$\mathbf{u} = \mathbf{N}\mathbf{u}_\mathrm{e} + f_\mathrm{e}\mathbf{N}\mathbf{a}_\mathrm{e} \tag{5}$$

$$\boldsymbol{\varepsilon} = \mathbf{B}\mathbf{u}_\mathrm{e} + f_\mathrm{e}\mathbf{B}\mathbf{a}_\mathrm{e} + (\boldsymbol{\nabla}_\varepsilon f_\mathrm{e})\ \mathbf{N}\mathbf{a}_\mathrm{e} \tag{6}$$

where $\mathbf{u}_\mathrm{e}$, $\mathbf{a}_\mathrm{e}$ are the element standard and enriched nodal variables and $\boldsymbol{\nabla}_\varepsilon$ is the symmetric gradient operator, so that $\mathbf{B}\,\mathbf{u}_\mathrm{e} = (\boldsymbol{\nabla}_\varepsilon \mathbf{N})\ \mathbf{u}_\mathrm{e}$.

Let $\mathbf{E}$ the elastic operator, such that the stress $\boldsymbol{\sigma}$ is given by $\boldsymbol{\sigma} = \mathbf{E}\boldsymbol{\varepsilon}$. The element internal virtual work is given by

$$\begin{aligned} L_\mathrm{i} = \int_{\Omega_\mathrm{e}} \boldsymbol{\varepsilon}^\mathrm{T}\boldsymbol{\sigma}\,\mathrm{d}\Omega = \\ \int_{\Omega_\mathrm{e}} \mathbf{B}^\mathrm{T}\mathbf{E}\,\mathbf{B}\,\mathrm{d}\Omega\,\mathbf{u}_\mathrm{e}\cdot\mathbf{u}_\mathrm{e} + \int_{\Omega_\mathrm{e}} \left(f_\mathrm{e}\mathbf{B}^\mathrm{T}\mathbf{E}\,\mathbf{B} + \mathbf{B}^\mathrm{T}\mathbf{E}\,(\boldsymbol{\nabla}_\varepsilon f_\mathrm{e})\,\mathbf{N}\right)\,\mathrm{d}\Omega\,\mathbf{a}_\mathrm{e}\cdot\mathbf{u}_\mathrm{e} + \\ \int_{\Omega_\mathrm{e}} \left(f_\mathrm{e}\mathbf{B}^\mathrm{T}\mathbf{E}\,\mathbf{B} + \mathbf{N}^\mathrm{T}\,(\boldsymbol{\nabla}_\varepsilon f_\mathrm{e})^\mathrm{T}\,\mathbf{E}\mathbf{B}\right)\,\mathrm{d}\Omega\,\mathbf{u}_\mathrm{e}\cdot\mathbf{a}_\mathrm{e} + \\ \int_{\Omega_\mathrm{e}} \Big(f_\mathrm{e}^2\mathbf{B}^\mathrm{T}\mathbf{E}\,\mathbf{B} + f_\mathrm{e}\mathbf{B}^\mathrm{T}\mathbf{E}\,(\boldsymbol{\nabla}_\varepsilon f_\mathrm{e})\,\mathbf{N} + f_\mathrm{e}\mathbf{N}^\mathrm{T}\,(\boldsymbol{\nabla}_\varepsilon f_\mathrm{e})^\mathrm{T}\,\mathbf{E}\mathbf{B} + \\ + \mathbf{N}^\mathrm{T}\,(\boldsymbol{\nabla}_\varepsilon f_\mathrm{e})^\mathrm{T}\,\mathbf{E}\,(\boldsymbol{\nabla}_\varepsilon f_\mathrm{e})\,\mathbf{N}\Big)\,\mathrm{d}\Omega\,\mathbf{a}_\mathrm{e}\cdot\mathbf{a}_\mathrm{e} \end{aligned} \tag{7}$$

The element stiffness matrix is then

$$\begin{aligned} \mathbf{K}_\mathrm{e} = \int_{\Omega_\mathrm{e}} \left[\begin{array}{c|c} \mathbf{k}_\mathrm{uu} & \mathbf{k}_\mathrm{ua} \\ \hline \mathbf{k}_\mathrm{au} & \mathbf{k}_\mathrm{aa} \end{array}\right] \mathrm{d}\Omega = \\ \int_{\Omega_\mathrm{e}} \left[\begin{array}{c|c} \mathbf{B}^\mathrm{T}\mathbf{E}\,\mathbf{B} & f_\mathrm{e}\mathbf{B}^\mathrm{T}\mathbf{E}\,\mathbf{B} + \mathbf{B}^\mathrm{T}\mathbf{E}\,(\boldsymbol{\nabla}_\varepsilon f_\mathrm{e})\,\mathbf{N} \\ \hline f_\mathrm{e}\mathbf{B}^\mathrm{T}\mathbf{E}\,\mathbf{B} + & f_\mathrm{e}^2\mathbf{B}^\mathrm{T}\mathbf{E}\,\mathbf{B} + f_\mathrm{e}\mathbf{B}^\mathrm{T}\mathbf{E}\,(\boldsymbol{\nabla}_\varepsilon f_\mathrm{e})\,\mathbf{N} + \\ \mathbf{N}^\mathrm{T}\,(\boldsymbol{\nabla}_\varepsilon f_\mathrm{e})^\mathrm{T}\,\mathbf{E}\mathbf{B} & f_\mathrm{e}\mathbf{N}^\mathrm{T}\,(\boldsymbol{\nabla}_\varepsilon f_\mathrm{e})^\mathrm{T}\,\mathbf{E}\mathbf{B} + \\ & \mathbf{N}^\mathrm{T}\,(\boldsymbol{\nabla}_\varepsilon f_\mathrm{e})^\mathrm{T}\,\mathbf{E}\,(\boldsymbol{\nabla}_\varepsilon f_\mathrm{e})\,\mathbf{N} \end{array}\right] \mathrm{d}\Omega \end{aligned} \tag{8}$$

where the matrix has been partitioned into four submatrices according to the ordering of nodal variables.

Note that, for discontinuous enrichment functions, the formal integral (8) on the element domain Ω_e needs to be split into two separate Gauss integrations, dividing the domain along the discontinuity surface Γ. For non-linear and non-polynomial enrichments high order Gauss rules or adaptive quadrature is needed for an accurate evaluation.

In the next sections it will be shown how the integrand function in (8) can be replaced by an equivalent polynomial expression defined in the whole domain Ω_e.

To this end, it will be assumed that the surface Γ will split the element domain Ω_e into two subdomains Ω_e^1 and Ω_e^2 such that

$$\Omega_e^1 \cup \Omega_e^2 = \Omega_e \quad ; \quad \Omega_e^1 \cap \Omega_e^2 = \Gamma \tag{9}$$

Although all the developments will carried out in the parent domain of the elements, when linear shape functions elements are considered these apply as well to the general case of isoparametric mapping, the Jacobian of the transformation being constant. In a more general setting the effect of isoparametric mapping has been outlined in [8].

3 Step Function Enrichment

Let consider the enrichment function $f_e = H$, Eq. (2), typically used in crack problems and suppose linear shape functions are used for the partition of unity [11]. It follows that function H is piecewise constant with zero derivative, Fig. 1a and:

- the strain matrix $\mathbf{B}$ is formed by constants for the linear bar, triangular and tetrahedral element;
- the strain matrix $\mathbf{B}$ is formed by linear functions for the bilinear quadrilateral element;
- $\boldsymbol{\nabla}_\varepsilon f_e$ is identically zero.

In the present case the element stiffness (8) simplifies to

$$\mathbf{K}_e = \int_{\Omega_e} \left[\begin{array}{c|c} \underline{\mathbf{B}^T\mathbf{E}\,\mathbf{B}} & \underline{H\mathbf{B}^T\mathbf{E}\,\mathbf{B}} \\ \hline H\mathbf{B}^T\mathbf{E}\,\mathbf{B} & H^2\mathbf{B}^T\mathbf{E}\,\mathbf{B} \end{array}\right] \mathrm{d}\Omega \tag{10}$$

where

$\mathbf{B}^T\mathbf{E}\,\mathbf{B}$ are continuous and differentiable functions;
$H\mathbf{B}^T\mathbf{E}\,\mathbf{B}$ are piecewise continuous and differentiable functions;
$H^2\mathbf{B}^T\mathbf{E}\,\mathbf{B}$ are continuous and differentiable functions ($H^2 = 1$ in Ω_e).

Therefore, only the off diagonal part $H\mathbf{B}^T\mathbf{E}\,\mathbf{B}$ is not Gauss integrable on the entire domain Ω_e, but is Gauss integrable on the two subdomains Ω_e^1 and Ω_e^2 so that the integration of (10) can be split as follows

$$\mathbf{K}_{\mathrm{e}} = \int_{\Omega_{\mathrm{e}}} \left[\begin{array}{c|c} \mathbf{B}^{\mathrm{T}}\mathbf{E}\,\mathbf{B} & \mathbf{0} \\ \hline \mathbf{0} & H^2\mathbf{B}^{\mathrm{T}}\mathbf{E}\,\mathbf{B} \end{array}\right] \mathrm{d}\Omega + \int_{\Omega_{\mathrm{e}}^1} H \left[\begin{array}{c|c} \mathbf{0} & \mathbf{B}^{\mathrm{T}}\mathbf{E}\,\mathbf{B} \\ \hline \mathbf{B}^{\mathrm{T}}\mathbf{E}\,\mathbf{B} & \mathbf{0} \end{array}\right] \mathrm{d}\Omega + \int_{\Omega_{\mathrm{e}}^2} H \left[\begin{array}{c|c} \mathbf{0} & \mathbf{B}^{\mathrm{T}}\mathbf{E}\,\mathbf{B} \\ \hline \mathbf{B}^{\mathrm{T}}\mathbf{E}\,\mathbf{B} & \mathbf{0} \end{array}\right] \mathrm{d}\Omega \quad (11)$$

We want to join the integration domains Ω_{e}^1 and Ω_{e}^2 by replacing the last two integrals in (11) with one defined on Ω_{e} giving the same numerical result. To this end let $\tilde{H}$ a polynomial defined by the formal equivalence

$$\int_{\Omega_{\mathrm{e}}} \tilde{H}\mathbf{B}^{\mathrm{T}}\mathbf{E}\,\mathbf{B}\,\mathrm{d}\Omega = \int_{\Omega_{\mathrm{e}}^1} H\mathbf{B}^{\mathrm{T}}\mathbf{E}\,\mathbf{B}\,\mathrm{d}\Omega + \int_{\Omega_{\mathrm{e}}^2} H\mathbf{B}^{\mathrm{T}}\mathbf{E}\,\mathbf{B}\,\mathrm{d}\Omega \quad (12)$$

If such polynomial $\tilde{H}$ exists, then it is, trivially

$$\mathbf{K}_{\mathrm{e}} = \int_{\Omega_{\mathrm{e}}} \left[\begin{array}{c|c} \mathbf{B}^{\mathrm{T}}\mathbf{E}\,\mathbf{B} & \tilde{H}\mathbf{B}^{\mathrm{T}}\mathbf{E}\,\mathbf{B} \\ \hline \tilde{H}\mathbf{B}^{\mathrm{T}}\mathbf{E}\,\mathbf{B} & H^2\mathbf{B}^{\mathrm{T}}\mathbf{E}\,\mathbf{B} \end{array}\right] \mathrm{d}\Omega \quad (13)$$

so that the integration can be performed on the entire element domain Ω_{e} with traditional Gauss quadrature.

The determination of the equivalent polynomial $\tilde{H}$ has been presented in [8] for one to three dimensional elements. The polynomial coefficients are functions of the signed distance function values at the element nodes.

4 Material Discontinuity Problems

The methodology introduced to deal with the jump in the displacement field due to cracks is applied to material discontinuity problems, i.e. to a discontinuity in the strain field. This kind of discontinuity can be represented by the enrichment function Q given by (3).

This case is slightly more involved than the previous one as the enrichment function has nonzero derivative and the elastic operator changes when crossing the discontinuity line.

Let $d(\mathbf{x}) \geq 0$ in Ω_e^1 and assume that the discontinuity line crosses completely an element. The enrichment function $Q(\mathrm{x}) = \mathrm{abs}(d(\mathrm{x}))$ has the property that its gradient is constant in Ω_e^1 and Ω_e^2 (the distance function being represented by a plane). It is

$$\boldsymbol{\nabla} Q(\mathbf{x_1}) = -\boldsymbol{\nabla} Q(\mathbf{x_2}) \quad \forall \mathbf{x_1} \in \Omega_e^1, \forall \mathbf{x_2} \in \Omega_e^2 \quad (14)$$

Consequently, introducing the signed distance function $H(\mathbf{x}) = \mathrm{sign}(d(\mathbf{x}))$, it is

$$\boldsymbol{\nabla} Q = H(\mathbf{x})\boldsymbol{\nabla} d \quad \forall \mathbf{x} \in \Omega_e \quad (15)$$

$$\boldsymbol{\nabla}_{\varepsilon} Q = H(\mathbf{x})\boldsymbol{\nabla}_{\varepsilon} d \quad \forall \mathbf{x} \in \Omega_e \quad (16)$$

Let

$$\boldsymbol{\nabla}_{\varepsilon} d = \boldsymbol{\mathcal{E}} \tag{17}$$

Replacing f_e by Q and $\boldsymbol{\nabla}_{\varepsilon} Q$ by $H\boldsymbol{\mathcal{E}}$, the element stiffness matrix (8) can be written

$$\mathbf{K}_{\mathrm{e}} = \int_{\Omega_{\mathrm{e}}} \left[\begin{array}{c|c} \mathbf{B}^{\mathrm{T}}\mathbf{E}\,\mathbf{B} & Q\mathbf{B}^{\mathrm{T}}\mathbf{E}\,\mathbf{B} + H\mathbf{B}^{\mathrm{T}}\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N} \\ \hline Q\mathbf{B}^{\mathrm{T}}\mathbf{E}\,\mathbf{B}+ & Q^2\mathbf{B}^{\mathrm{T}}\mathbf{E}\,\mathbf{B} + Q\,H\mathbf{B}^{\mathrm{T}}\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N}+ \\ H\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\mathbf{E}\,\mathbf{B} & Q\,H\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\mathbf{E}\,\mathbf{B} + H^2\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N} \end{array}\right] \mathrm{d}\Omega \tag{18}$$

Suppose the two regions Ω_e^1 and Ω_e^2, separated by the discontinuity line, have elastic operators $\mathbf{E}_1$ and $\mathbf{E}_2$, respectively. Let

$$\mathbf{E}_{\mathrm{m}} = \frac{1}{2}\left(\mathbf{E}_1 + \mathbf{E}_2\right) \tag{19}$$

$$\Delta\mathbf{E} = \frac{1}{2}\left(\mathbf{E}_1 - \mathbf{E}_2\right) \tag{20}$$

Then the elastic operator can be written as

$$\mathbf{E} = \mathbf{E}_{\mathrm{m}} + H(\mathbf{x})\,\Delta\mathbf{E} \tag{21}$$

so that (18) is split into the sum of two parts

$$\begin{aligned} \mathbf{K}_{\mathrm{e}} = &\int_{\Omega_{\mathrm{e}}} \left[\begin{array}{c|c} \mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B} & Q\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B} + H\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\boldsymbol{\mathcal{E}}\mathbf{N} \\ \hline Q\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B}+ & Q^2\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B} + Q\,H\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\boldsymbol{\mathcal{E}}\mathbf{N}+ \\ H\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B} & Q\,H\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\mathbf{B} + H^2\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\boldsymbol{\mathcal{E}}\mathbf{N} \end{array}\right] \mathrm{d}\Omega + \\ &+ \int_{\Omega_{\mathrm{e}}} \left[\begin{array}{c|c} H\mathbf{B}^{\mathrm{T}}\Delta\mathbf{E}\,\mathbf{B} & H\,Q\mathbf{B}^{\mathrm{T}}\Delta\mathbf{E}\,\mathbf{B} + H^2\mathbf{B}^{\mathrm{T}}\Delta\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N} \\ \hline H\,Q\mathbf{B}^{\mathrm{T}}\Delta\mathbf{E}\,\mathbf{B}+ & H\,Q^2\mathbf{B}^{\mathrm{T}}\Delta\mathbf{E}\,\mathbf{B} + H^2Q\mathbf{B}^{\mathrm{T}}\Delta\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N}+ \\ H^2\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\Delta\mathbf{E}\,\mathbf{B} & H^2\,Q\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\Delta\mathbf{E}\,\mathbf{B} + H^3\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\Delta\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N} \end{array}\right] \mathrm{d}\Omega \end{aligned} \tag{22}$$

With the above transformations, all the discontinuities have been concentrated into the functions Q and H, so that (22) can be split into continuous and differentiable terms, and piecewise continuous and differentiable terms. Table 1 summarizes the properties of the integrand functions.

To evidence continuous and differentiable terms in Ω_e and continuous and differentiable terms in the subdomains Ω_e^1 and Ω_e^2, (22) is written

Table 1. Continuity of the integrand functions for material discontinuity.

Continuous and differentiable	Piecewise continuous and differentiable
H^2	H
Q^2	Q
$H\,Q$	H^3
	$H\,Q^2$
	$H^2\,Q$

$$\mathbf{K}_{\mathrm{e}} = \int_{\Omega_{\mathrm{e}}} \left[\begin{array}{c|c} \mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B} & \mathbf{0} \\ \hline \mathbf{0} & \begin{array}{l} Q^2\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B} + Q\,H\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\boldsymbol{\mathcal{E}}\mathbf{N}+ \\ Q\,H\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\mathbf{B} + H^2\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\boldsymbol{\mathcal{E}}\mathbf{N} \end{array} \end{array}\right] \mathrm{d}\Omega+$$

$$+ \int_{\Omega_{\mathrm{e}}^1\cup\Omega_{\mathrm{e}}^2} \left[\begin{array}{c|c} \mathbf{0} & \begin{array}{l} Q\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B}+ \\ H\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\boldsymbol{\mathcal{E}}\mathbf{N} \end{array} \\ \hline \begin{array}{l} Q\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B}+ \\ H\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B} \end{array} & \mathbf{0} \end{array}\right] \mathrm{d}\Omega+$$

$$+ \int_{\Omega_{\mathrm{e}}} \left[\begin{array}{c|c} \mathbf{0} & \begin{array}{l} H\,Q\mathbf{B}^{\mathrm{T}}\boldsymbol{\Delta}\mathbf{E}\,\mathbf{B}+ \\ H^2\mathbf{B}^{\mathrm{T}}\boldsymbol{\Delta}\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N} \end{array} \\ \hline \begin{array}{l} H\,Q\mathbf{B}^{\mathrm{T}}\boldsymbol{\Delta}\mathbf{E}\,\mathbf{B}+ \\ H^2\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\boldsymbol{\Delta}\mathbf{E}\,\mathbf{B} \end{array} & \mathbf{0} \end{array}\right] \mathrm{d}\Omega+$$

$$+ \int_{\Omega_{\mathrm{e}}^1\cup\Omega_{\mathrm{e}}^2} \left[\begin{array}{c|c} H\mathbf{B}^{\mathrm{T}}\boldsymbol{\Delta}\mathbf{E}\,\mathbf{B} & \mathbf{0} \\ \hline \mathbf{0} & \begin{array}{l} H\,Q^2\mathbf{B}^{\mathrm{T}}\boldsymbol{\Delta}\mathbf{E}\,\mathbf{B} + H^2Q\mathbf{B}^{\mathrm{T}}\boldsymbol{\Delta}\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N}+ \\ H^2\,Q\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\boldsymbol{\Delta}\mathbf{E}\,\mathbf{B} + H^3\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\boldsymbol{\Delta}\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N} \end{array} \end{array}\right] \mathrm{d}\Omega \quad (23)$$

and, grouping together the terms with the same continuity and differentiability properties,

$$\mathbf{K}_{\mathrm{e}} = \int_{\Omega_{\mathrm{e}}} \left[\begin{array}{c|c} \mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B} & H\,Q\mathbf{B}^{\mathrm{T}}\boldsymbol{\Delta}\mathbf{E}\,\mathbf{B} + H^2\mathbf{B}^{\mathrm{T}}\boldsymbol{\Delta}\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N} \\ \hline \begin{array}{l} H\,Q\mathbf{B}^{\mathrm{T}}\boldsymbol{\Delta}\mathbf{E}\,\mathbf{B}+ \\ H^2\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\boldsymbol{\Delta}\mathbf{E}\,\mathbf{B} \end{array} & \begin{array}{l} Q^2\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B} + Q\,H\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\boldsymbol{\mathcal{E}}\mathbf{N}+ \\ Q\,H\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\mathbf{B} + H^2\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\boldsymbol{\mathcal{E}}\mathbf{N} \end{array} \end{array}\right] \mathrm{d}\Omega+$$

$$+ \int_{\Omega_{\mathrm{e}}^1\cup\Omega_{\mathrm{e}}^2} \left[\begin{array}{c|c} H\mathbf{B}^{\mathrm{T}}\boldsymbol{\Delta}\mathbf{E}\,\mathbf{B} & Q\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B} + H\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\boldsymbol{\mathcal{E}}\mathbf{N} \\ \hline \begin{array}{l} Q\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B}+ \\ H\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B} \end{array} & \begin{array}{l} H\,Q^2\mathbf{B}^{\mathrm{T}}\boldsymbol{\Delta}\mathbf{E}\,\mathbf{B} + H^2Q\mathbf{B}^{\mathrm{T}}\boldsymbol{\Delta}\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N}+ \\ H^2Q\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\boldsymbol{\Delta}\mathbf{E}\,\mathbf{B} + H^3\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\boldsymbol{\Delta}\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N} \end{array} \end{array}\right] \mathrm{d}\Omega \quad (24)$$

The first term on the r.h.s. of (24) can be numerically evaluated by Gauss quadrature, while the second should be evaluated separately on the two subdomains Ω_e^1 and Ω_e^2. In the light of the proposed methodology two equivalent polynomials $\tilde{H}$ and $\tilde{Q}$ are to be determined as a replacement to the discontinuous functions, satisfying the following equation

$$\int_{\Omega_{\mathrm{e}}^1} \left[\begin{array}{c|c} H\mathbf{B}^{\mathrm{T}}\Delta\mathbf{E}\,\mathbf{B} & Q\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B} + H\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\boldsymbol{\mathcal{E}}\mathbf{N} \\ \hline Q\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B}+ & H\,Q^2\mathbf{B}^{\mathrm{T}}\Delta\mathbf{E}\,\mathbf{B} + H^2Q\mathbf{B}^{\mathrm{T}}\Delta\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N}+ \\ H\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B} & H^2\,Q\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\Delta\mathbf{E}\,\mathbf{B} + H^3\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\Delta\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N} \end{array}\right] \mathrm{d}\Omega +$$

$$+ \int_{\Omega_{\mathrm{e}}^2} \left[\begin{array}{c|c} H\mathbf{B}^{\mathrm{T}}\Delta\mathbf{E}\,\mathbf{B} & Q\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B} + H\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\boldsymbol{\mathcal{E}}\mathbf{N} \\ \hline Q\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B}+ & H\,Q^2\mathbf{B}^{\mathrm{T}}\Delta\mathbf{E}\,\mathbf{B} + H^2Q\mathbf{B}^{\mathrm{T}}\Delta\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N}+ \\ H\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B} & H^2\,Q\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\Delta\mathbf{E}\,\mathbf{B} + H^3\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\Delta\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N} \end{array}\right] \mathrm{d}\Omega =$$

$$= \int_{\Omega_{\mathrm{e}}} \left[\begin{array}{c|c} \tilde{H}\mathbf{B}^{\mathrm{T}}\Delta\mathbf{E}\,\mathbf{B} & \tilde{Q}\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B} + \tilde{H}\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\boldsymbol{\mathcal{E}}\mathbf{N} \\ \hline \tilde{Q}\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B}+ & \tilde{H}\,Q^2\mathbf{B}^{\mathrm{T}}\Delta\mathbf{E}\,\mathbf{B} + H^2\tilde{Q}\mathbf{B}^{\mathrm{T}}\Delta\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N}+ \\ \tilde{H}\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B} & H^2\,\tilde{Q}\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\Delta\mathbf{E}\,\mathbf{B} + \tilde{H}\,H^2\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\Delta\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N} \end{array}\right] \mathrm{d}\Omega \quad (25)$$

The problem of the determination of $\tilde{Q}$ and $\tilde{H}$ has been examined and solved with reference to 1D and 2D problems in [8]. Tables 2 and 3 summarize the polynomial degree of the terms multiplying $\tilde{H}$ and $\tilde{Q}$ in (25).

Table 2. Polynomial degree of the terms multiplying $\tilde{H}$ in (25).

Term	Degree
$\mathbf{B}^{\mathrm{T}}\Delta\mathbf{E}\,\mathbf{B}$	quadratic
$\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\boldsymbol{\mathcal{E}}\mathbf{N}$	cubic (incomplete)
$\mathbf{N}^{\mathrm{T}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}\Delta\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N}$	quartic (incomplete)
$Q^2\mathbf{B}^{\mathrm{T}}\Delta\mathbf{E}\,\mathbf{B}$	quartic

Table 3. Polynomial degree of the terms multiplying $\tilde{Q}$ in (25).

Term	Degree
$\mathbf{B}^{\mathrm{T}}\mathbf{E}_{\mathrm{m}}\,\mathbf{B}$	quadratic
$\mathbf{B}^{\mathrm{T}}\Delta\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N}$	cubic (incomplete)

5 Regularized Heaviside Function

In the present section is examined the case the enrichment function is given by the regularized Heaviside function $f_{\mathrm{e}} = H_\rho$, Eq. (4). By the chain rule and (17) it is

$$\boldsymbol{\nabla}_\varepsilon H_\rho = \frac{\partial H_\rho}{\partial d}\,\boldsymbol{\mathcal{E}} = \delta\,\boldsymbol{\mathcal{E}} \quad (26)$$

where it has been set $\delta = \partial H_\rho / \partial d$. The element stiffness (8) in this particular case is therefore

$$\mathbf{K}_\mathrm{e} = \int_{\Omega_\mathrm{e}} \left[\begin{array}{c|c} \mathbf{B}^\mathrm{T}\mathbf{E}\,\mathbf{B} & H_\rho \mathbf{B}^\mathrm{T}\mathbf{E}\,\mathbf{B} + \delta \mathbf{B}^\mathrm{T}\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N} \\ \hline H_\rho \mathbf{B}^\mathrm{T}\mathbf{E}\,\mathbf{B} + & H_\rho^2 \mathbf{B}^\mathrm{T}\mathbf{E}\,\mathbf{B} + \delta H_\rho \mathbf{B}^\mathrm{T}\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N} + \\ \delta \mathbf{N}^\mathrm{T}\boldsymbol{\mathcal{E}}^\mathrm{T}\mathbf{E}\mathbf{B} & \delta H_\rho \mathbf{N}^\mathrm{T}\boldsymbol{\mathcal{E}}^\mathrm{T}\mathbf{E}\mathbf{B} + \delta^2 \mathbf{N}^\mathrm{T}\boldsymbol{\mathcal{E}}^\mathrm{T}\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N} \end{array}\right] \mathrm{d}\Omega \tag{27}$$

where $\boldsymbol{\mathcal{E}}$ is a constant tensor. The examination of (27) shows that the non-polynomial terms are

- H_ρ, multiplying the constants $\mathbf{B}^\mathrm{T}\mathbf{E}\,\mathbf{B}$;
- δ, multiplying the first order polynomials $\mathbf{B}^\mathrm{T}\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N}$;
- H_ρ^2, multiplying the constants $\mathbf{B}^\mathrm{T}\mathbf{E}\,\mathbf{B}$;
- δH_ρ, multiplying the first order polynomials $\mathbf{B}^\mathrm{T}\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N}$;
- δ^2, multiplying the second order polynomials $\mathbf{N}^\mathrm{T}\boldsymbol{\mathcal{E}}^\mathrm{T}\mathbf{E}\boldsymbol{\mathcal{E}}\mathbf{N}$.

For each of these non-polynomial functions an equivalent polynomial can be defined for replacement in (27), following the path introduced in [8]. As previously pointed out, the equivalence is set so that the integral of the polynomial is equal to the integral of the original function in each finite element Ω_e.

The replacement of the original functions with the equivalent polynomials allows to use traditional low order Gauss quadrature without introducing any approximation in the evaluation of the element stiffness.

To illustrate the procedure, reference is made to the one-dimensional case, ξ being the abscissa in the parent element domain $[-1, +1]$.

- As function H_ρ multiplies constants, its equivalent polynomial $\mathcal{P}_H$ must have the property

$$\int_{-1}^{+1} H_\rho \,\mathrm{d}\xi = \int_{-1}^{+1} \mathcal{P}_H \,\mathrm{d}\xi \tag{28}$$

- As function δ multiplies first order polynomials, its equivalent polynomial $\mathcal{P}_\delta$ must have the property

$$\int_{-1}^{+1} \delta \,\mathrm{d}\xi = \int_{-1}^{+1} \mathcal{P}_\delta \,\mathrm{d}\xi \tag{29a}$$

$$\int_{-1}^{+1} \delta\, \xi \,\mathrm{d}\xi = \int_{-1}^{+1} \mathcal{P}_\delta\, \xi \,\mathrm{d}\xi \tag{29b}$$

- As function H_ρ^2 multiplies constants, its equivalent polynomial $\mathcal{P}_{H^2}$ must have the property

$$\int_{-1}^{+1} H_\rho^2 \,\mathrm{d}\xi = \int_{-1}^{+1} \mathcal{P}_{H^2} \,\mathrm{d}\xi \tag{30}$$

- As function δH_ρ multiplies first order polynomials, its equivalent polynomial $\mathcal{P}_{\delta H}$ must have the property

$$\int_{-1}^{+1} \delta\, H_\rho \,\mathrm{d}\xi = \int_{-1}^{+1} \mathcal{P}_{\delta H} \,\mathrm{d}\xi \tag{31a}$$

$$\int_{-1}^{+1} \delta\, H_\rho\, \xi \,\mathrm{d}\xi = \int_{-1}^{+1} \mathcal{P}_{\delta H}\, \xi \,\mathrm{d}\xi \tag{31b}$$

- As function δ^2 multiplies second order polynomials, its equivalent polynomial $\mathcal{P}_{\delta^2}$ must have the property

$$\int_{-1}^{+1} \delta^2 \,\mathrm{d}\xi = \int_{-1}^{+1} \mathcal{P}_{\delta^2} \,\mathrm{d}\xi \tag{32a}$$

$$\int_{-1}^{+1} \delta^2\, \xi \,\mathrm{d}\xi = \int_{-1}^{+1} \mathcal{P}_{\delta^2}\, \xi \,\mathrm{d}\xi \tag{32b}$$

$$\int_{-1}^{+1} \delta^2\, \xi^2 \,\mathrm{d}\xi = \int_{-1}^{+1} \mathcal{P}_{\delta^2}\, \xi^2 \,\mathrm{d}\xi \tag{32c}$$

The left hand sides of Equations (28)…(32) can be analytically evaluated for the assumed function H_ρ, Eq. (4). Assuming $\mathcal{P}_H$ and $\mathcal{P}_{H^2}$ as constants, $\mathcal{P}_\delta$ and $\mathcal{P}_{\delta H}$ as first order polynomials and $\mathcal{P}_{\delta^2}$ as second order polynomial, it is straightforward to observe that Equations (28)…(32) are linear systems in the unknown polynomial coefficients of the form

$$\mathbf{A}\,\mathbf{c} = \mathbf{b} \tag{33}$$

where $\mathbf{c}$ is the column vector of the variables (polynomial coefficients) and $\mathbf{b}$ is the column vector of solutions.

The vectors of solutions are the right end sides of the equations, listing the analytic integrals on the element domain and will be function of the position of the surface Γ, H_ρ being function of the signed distance d.

The determination of the equivalent polynomials has been carried out for the linear bar and triangle. For the sake of brevity, in the Appendix to the present work the solution will be given for the one-dimensional case, being the 2D case long to list.

6 Gaussian Quadrature for the Equivalent Polynomials

In the previous sections it has been shown that, replacing the enrichment function by equivalent polynomials allows for the evaluation of the stiffness matrix by Gauss quadrature without splitting the element domain.

The results are summarized in Table 4, for the examined problems and finite elements. As evident from the Table, in the case of cracks and material

discontinuity the proposed method avoids the quadrature domain splitting at the cost of sometimes doubling the polynomial degree of the integrand functions. For crack problems with triangular or tetrahedral elements the degree of the integrand remains the same, yielding an additional advantage.

For the regularized Heaviside function a numerically exact Gauss quadrature rule does not exist and comments and results are illustrated in the next section.

Table 4. Maximum polynomial degree of the integrand functions for the examined problems (element parent domain).

Problem	Element	Max polynomial degree (quadrature on subcells)	Equivalent polynomial	Max equivalent polynomial degree (quadrature on Ω_e)
crack	2 nodes bar, 3 nodes triangle, 4 nodes tetrahedron	0	$\tilde{H}$	0
	4 nodes quadrilateral	2	$\tilde{H}$	4
material discont.	2 nodes bar, 3 nodes triangle, 4 nodes tetrahedron	2	$\tilde{H}$, $\tilde{Q}$	4
	4 nodes quadrilateral	4	$\tilde{H}$, $\tilde{Q}$	8
regularized Heaviside	2 nodes bar, 3 nodes triangle, 4 nodes tetrahedron	n.a.	$\mathcal{P}_H$, $\mathcal{P}_{H^2}$	0
		n.a.	$\mathcal{P}_\delta$, $\mathcal{P}_{\delta H}$	1
		n.a.	$\mathcal{P}_{\delta^2}$	2

6.1 Quadrature errors

In this section attention will be focused on the quadrature of the regularized Heaviside step function (4). This function and its derivative multiply from constants to second degree polynomials, see Table 5 and is to be integrated in the element domain. In this analysis the one dimensional case is examined as the higher dimensional cases are expected to behave similarly.

To study the influence of the quadrature method in the evaluation of the terms, the errors in the integral of the function times the monomial terms, as reported in Table 5, have been computed.

Table 5. Multiplying monomials for the enrichment function and derivative.

Function	Multiplying monomials
H_ρ	1
δ	1, ξ
H_ρ^2	1
δH_ρ	1, ξ
δ^2	1, ξ, ξ^2

Three integration methods are considered:

- analytic integration, providing the reference exact value;
- equivalent polynomial integration, giving the same exact values as analytic integration. These results will not be listed being equal to the above;
- Gauss integration on the entire element domain;
- Gauss integration by splitting the element in two integration subdomains at the abscissa where $d = 0$;

The results are reported in tables where the position of the "discontinuity" $d = 0$ has been fixed at the abscissa $\xi_d = 0.5$, then two values of ρ, $\rho = 0.1$ and $\rho = 0.001$, and four Gauss rules, 2, 5, 10, 50 Gauss points have been considered. All the computations have been performed developing a software with Mathematica® by Wolfram Research.

For each quadrature term Tables 6 ... 12 report the exact value of the integral and the percentage error by using the Gauss rule over the entire element domain (1 dom.) and by splitting the element into two subdomains separated by the line $d = 0$, i.e. at $\xi_d = 0.5$, and applying the rule in each subdomain (2 dom.). Therefore, the two domain quadrature actually uses two times the number of Gauss points compared to the quadrature over the entire domain.

An important consideration is to be made for a proper reading of the results: as some of the integrals have zero as exact limit value when the regularization parameter $\rho \to 0$, the error in this case has basically no meaning, being influenced by truncation errors to a large extent.

The two examined cases, $\rho = 0.1$ and $\rho = 0.001$, are such that in the first the function varies rather smoothly in the domain and its first derivative δ is numerically nonvanishing; on the other hand, in the second case, H_ρ numerically approximates quite closely the Heaviside step function and, numerically, it presents a vanishing derivative except at the point $\xi = \xi_d$ where it attains a very large value.

Assume the meaningful error in the evaluation of the stiffness is due to the terms whose absolute integral value is greater than 1% the largest integral. With reference to Table 6 the largest absolute integral value is 1 and this means the only meaningful terms are on the first four rows, H_ρ, H_ρ^2, δ and $\delta\xi$.

Examine the case $\rho = 0.1$, Tables 6 . . . 8. The error decreases as the number of Gauss points is increased, as expected. Although the regularized Heaviside function is everywhere continuous and differentiable, we observe that splitting the domain into two subdomains yields much better results than the single domain quadrature. The single domain quadrature gives satisfactory results only with the 10 points rule, that is an excessive number, especially if we think of higher dimensional problems (2D and 3D).

If reference is made to the case $\rho = 0.001$, Tables 9 . . . 12, we observe again that domain splitting is necessary. In fact, if two domains are considered, the 5 points Gauss rule gives satisfactory results, while single domain quadrature requires a 50 Gauss points rule to give an error less than 5% in the largest integral.

In general, it may be concluded that the regularized Heaviside function requires splitting of the integration domain and a high order quadrature rule for a proper evaluation of the stiffness terms. Unless very high order rules are adopted, single domain Gauss quadrature results in large errors.

Comparing these results to the polynomial mapping technique we observe that, although the evaluation of the polynomial coefficients is slightly time consuming (see the expressions of the coefficients in the Appendix), the exact value of each term is obtained with a very low order Gauss rule, see Table 4. Moreover, no splitting of the domain is required, with the consequent simplifications in 2D and, especially, 3D problems.

Table 6. Quadrature results for $\rho = 0.1$ and the 2 Gauss points quadrature. The equivalent polynomial results are not listed, being exact. The error of classical Gauss quadrature on the entire domain and by domain splitting are reported in the last two columns.

Function	Exact value of integral	% of error in Gauss quadrature, 1 dom.	% of error in Gauss quadrature, 2 dom.
H_ρ	−0.99933	−53.8 %	+6.1 %
H_ρ^2	+1.70135	−24.2 %	+4.9 %
δ	+0.01993	+131.5 %	−38.2 %
$\delta\,\xi$	+0.00993	+168.4 %	−36.8 %
$\delta\,H_\rho$	−0.00007	−37105 %	−4755 %
$\delta\,H_\rho\,\xi$	+0.00143	+906.2 %	+131.9 %
δ^2	+0.00010	+112.9 %	−68.4 %
$\delta^2\,\xi$	+0.00050	+145.8 %	−62.8 %
$\delta^2\,\xi^2$	+0.00025	+178.3 %	−56.1 %

Table 7. Quadrature results for $\rho = 0.1$ and the 5 Gauss points quadrature. The equivalent polynomial results are not listed, being exact. The error of classical Gauss quadrature on the entire domain and by domain splitting are reported in the last two columns.

Function	Exact value of integral	% of error in Gauss quadrature, 1 dom.	% of error in Gauss quadrature, 2 dom.
H_ρ	−0.99933	−10.4 %	+0.1 %
H_ρ^2	+1.70135	−8.6 %	−0.3 %
δ	+0.01993	+67.4 %	−0.3 %
$\delta\,\xi$	+0.00993	+80.5 %	−0.6 %
$\delta\,H_\rho$	−0.00007	−15621 %	+826 %
$\delta\,H_\rho\,\xi$	+0.00143	+318.3 %	−22.4 %
δ^2	+0.00010	+121.8 %	−6.2 %
$\delta^2\,\xi$	+0.00050	+138.9 %	−7.5 %
$\delta^2\,\xi^2$	+0.00025	+152.4 %	−8.7 %

Table 8. Quadrature results for $\rho = 0.1$ and the 10 Gauss points quadrature. The equivalent polynomial results are not listed, being exact. The error of classical Gauss quadrature on the entire domain and by domain splitting are reported in the last two columns.

Function	Exact value of integral	% of error in Gauss quadrature, 1 dom.	% of error in Gauss quadrature, 2 dom.
H_ρ	−0.99933	+2.4 %	+0.0 %
H_ρ^2	+1.70135	−0.1 %	+0.0 %
δ	+0.01993	−5.4 %	+0.0 %
$\delta\,\xi$	+0.00993	−9.5 %	+0.0 %
$\delta\,H_\rho$	−0.00007	+6091 %	+0.2 %
$\delta\,H_\rho\,\xi$	+0.00143	−141.2 %	−0.0 %
δ^2	+0.00010	−22.5 %	−0.0 %
$\delta^2\,\xi$	+0.00050	−29.9 %	−0.0 %
$\delta^2\,\xi^2$	+0.00025	−36.3 %	−0.0 %

7 Conclusions

In the paper a method for eliminating the introduction of quadrature subcells when using discontinuous/non-differentiable or highly non-linear enrichment functions in local Partition of Unity methods has been introduced. The method replaces the enrichment functions with equivalent polynomials, defined to yield the same value of the original enrichment function in the integration of the stiffness coefficients. Consequently, the employment of the equivalent polynomials allows low order Gauss quadrature to be be applied to the entire element domain. The method is deeply tied to the level set rep-

Table 9. Quadrature results for $\rho = 0.001$ and the 2 Gauss points quadrature. The equivalent polynomial results are not listed, being exact. The error of classical Gauss quadrature on the entire domain and by domain splitting are reported in the last two columns.

Function	Exact value of integral	% of error in Gauss quadrature, 1 dom.	% of error in Gauss quadrature, 2 dom.
H_ρ	−1.00000	−100.0 %	+0.0 %
H_ρ^2	+1.99700	+0.2 %	+0.2 %
δ	+0.00000	−100.0 %	−100.0 %
$\delta\,\xi$	+0.00000	−100.0 %	−100.0 %
$\delta\,H_\rho$	−0.00000	−100.0 %	−100.0 %
$\delta\,H_\rho\,\xi$	+0.00000	−100.0 %	−100.0 %
δ^2	+0.00000	−100.0 %	−100.0 %
$\delta^2\,\xi$	+0.00000	−100.0 %	−100.0 %
$\delta^2\,\xi^2$	+0.00000	−100.0 %	−100.0 %

Table 10. Quadrature results for $\rho = 0.001$ and the 5 Gauss points quadrature. The equivalent polynomial results are not listed, being exact. The error of classical Gauss quadrature on the entire domain and by domain splitting are reported in the last two columns.

Function	Exact value of integral	% of error in Gauss quadrature, 1 dom.	% of error in Gauss quadrature, 2 dom.
H_ρ	−1.00000	−43.1 %	+0.0 %
H_ρ^2	+1.99700	+0.2 %	+0.2 %
δ	+0.00000	−100.0 %	−100.0 %
$\delta\,\xi$	+0.00000	−100.0 %	−100.0 %
$\delta\,H_\rho$	−0.00000	−82.4 %	+0.0 %
$\delta\,H_\rho\,\xi$	+0.00000	−100.0 %	−100.0 %
δ^2	+0.00000	−100.0 %	−100.0 %
$\delta^2\,\xi$	+0.00000	−100.0 %	−100.0 %
$\delta^2\,\xi^2$	+0.00000	−100.0 %	−100.0 %

resentation of the discontinuity surface, as the data defining the discontinuity are used to generate the closed form solution of the equivalent polynomial. The method is illustrated with reference to linear shape function elements for crack, material discontinuity and regularized jump problems in one, two and three dimensions. In particular a regularized Heaviside function is considered and the polynomials are given in the one dimensional case. An extensive error analysis has been performed in this case, showing domain splitting along the discontinuity surface is always necessary to properly evaluate the stiffness terms.

Table 11. Quadrature results for $\rho = 0.001$ and the 10 Gauss points quadrature. The equivalent polynomial results are not listed, being exact. The error of classical Gauss quadrature on the entire domain and by domain splitting are reported in the last two columns.

Function	Exact value of integral	% of error in Gauss quadrature, 1 dom.	% of error in Gauss quadrature, 2 dom.
H_ρ	-1.00000	$+13.0$ %	$+0.0$ %
H_ρ^2	$+1.99700$	$+0.2$ %	$+0.1$ %
δ	$+0.00000$	-100.0 %	-98.8 %
$\delta\,\xi$	$+0.00000$	-100.0 %	-98.8 %
$\delta\,H_\rho$	-0.00000	-100.0 %	$+0.0$ %
$\delta\,H_\rho\,\xi$	$+0.00000$	-100.0 %	$+725.0$ %
δ^2	$+0.00000$	-100.0 %	-100.0 %
$\delta^2\,\xi$	$+0.00000$	-100.0 %	-100.0 %
$\delta^2\,\xi^2$	$+0.00000$	-100.0 %	-100.0 %

Table 12. Quadrature results for $\rho = 0.001$ and the 50 Gauss points quadrature. The equivalent polynomial results are not listed, being exact. The error of classical Gauss quadrature on the entire domain and by domain splitting are reported in the last two columns.

Function	Exact value of integral	% of error in Gauss quadrature, 1 dom.	% of error in Gauss quadrature, 2 dom.
H_ρ	-1.00000	-4.5 %	$+0.0$ %
H_ρ^2	$+1.99700$	$+0.1$ %	$+0.1$ %
δ	$+0.00000$	-68.9 %	-98.8 %
$\delta\,\xi$	$+0.00000$	-68.9 %	-98.8 %
$\delta\,H_\rho$	-0.00000	$+10^{15}$ %	$+0.000$ %
$\delta\,H_\rho\,\xi$	$+0.00000$	$+20582$ %	$+725$ %
δ^2	$+0.00000$	-99.3 %	-100.0 %
$\delta^2\,\xi$	$+0.00000$	-99.3 %	-100.0 %
$\delta^2\,\xi^2$	$+0.00000$	-99.3 %	-100.0 %

Acknowledgments

The support of the Italian Ministry of the University and Scientific Research under program PRIN 2004 – Nanotribology is gratefully acknowledged.

Appendix

Equivalent polynomials for the 1D bar in the parent domain $\xi \in [-1, 1]$.

$$\mathcal{P}_H = \frac{-2\,e^{\frac{1+\xi_d}{\rho}}\,\xi_d - \rho + e^{\frac{2\,\xi_d}{\rho}}\,\rho}{2\,e^{\frac{1+\,\xi_d}{\rho}}}$$

$$\mathcal{P}_{H^2} = \frac{e^{\frac{2\,(1+\xi_d)}{\rho}}\,(4-6\,\rho) - \rho - e^{\frac{4\,\xi_d}{\rho}}\,\rho + 4\,e^{\frac{1+\xi_d}{\rho}}\,\rho + 4\,e^{\frac{1+3\,\xi_d}{\rho}}\,\rho}{4\,e^{\frac{2\,(1+\xi_d)}{\rho}}}$$

$$\mathcal{P}_\delta = -\frac{\left(1+e^{\frac{2\,\xi_d}{\rho}} - 2\,e^{\frac{1+\xi_d}{\rho}}\right)\,\rho^{\,2}}{2\,e^{\frac{1+\xi_d}{\rho}}} + \\ + \xi\,\frac{-3\,\rho^2\,\left(-1-2\,e^{\frac{1+\xi_d}{\rho}}\,\xi_d - \rho + e^{\frac{2\,\xi_d}{\rho}}\,(1+\rho)\right)}{2\,e^{\frac{1+\,\xi_d}{\rho}}}$$

$$\mathcal{P}_{\delta H} = \frac{\left(-1+e^{\frac{\xi_d}{\rho}}\right)\,\left(1+e^{\frac{\xi_d}{\rho}}\right)\,\left(1+e^{\frac{2\,\xi_d}{\rho}} - 2\,e^{\frac{1+\xi_d}{\rho}}\right)\,\rho^{\,2}}{4\,e^{\frac{2\,(1+\xi_d)}{\rho}}} + \\ + \xi\,\frac{3\,\rho^2}{8\,e^{\frac{2\,(1+\xi_d)}{\rho}}}\,\left(2+\rho+6\,e^{\frac{2\,(1+\xi_d)}{\rho}}\,\rho + \right. \\ \left. -4\,e^{\frac{1+\xi_d}{\rho}}\,(1+\rho) - 4\,e^{\frac{1+3\,\xi_d}{\rho}}\,(1+\rho) + e^{\frac{4\,\xi_d}{\rho}}\,(2+\rho)\right)$$

$$\mathcal{P}_{\delta^2} = \frac{3\,\rho^3}{32\,e^{\frac{2\,(1+\xi_d)}{\rho}}}\,\left(4+10\,\rho+5\,\rho^2 - 2\,e^{\frac{2\,(1+\xi_d)}{\rho}}\,\left(-6+10\,{\xi_d}^2+5\,\rho^2\right) + \right. \\ \left. + e^{\frac{4\,\xi_d}{\rho}}\,\left(4+10\,\rho+5\,\rho^2\right)\right) + \\ - \xi\,\frac{3\,\rho^3\,\left(-2-4\,e^{\frac{2\,(1+\xi_d)}{\rho}}\,\xi_d - \rho + e^{\frac{4\,\xi_d}{\rho}}\,(2+\rho)\right)}{8\,e^{\frac{2\,(1+\xi_d)}{\rho}}} + \\ - \xi^2\,\frac{15\,\rho^3}{32\,e^{\frac{2\,(1+\xi_d)}{\rho}}}\,\left(4+6\,\rho+3\,\rho^2 - 2\,e^{\frac{2\,(1+\xi_d)}{\rho}}\,\left(-2+6\,{\xi_d}^2+3\,\rho^2\right) + \right. \\ \left. + e^{\frac{4\,\xi_d}{\rho}}\,\left(4+6\,\rho+3\,\rho^2\right)\right)$$

References

1. Melenk J M, Babuska I (1996) Comp Meth Appl Mech Engng 39:289–314.
2. Belytschko T, Black T (1999) Int J Numer Meth Engng 45:601–620.
3. Belytschko T, Moës N, Usui S, Parimi C (2001) Int J Numer Meth Engng 50:993–1013.
4. Sukumar N, Chopp D L, Moës N, Belytschko T (2001) Comp Meth Appl Mech Engng 90:6183–6200.
5. Chessa J, Smolinski P, Belytschko T (2002) Int J Numer Meth Engng 53:1957–1977.
6. Moës N, Gravouil A, Belytschko T (2002) Int J Numer Meth Engng 53:2549–2568.

7. Sukumar N, Moës N, Moran B, Belytschko T (2000) Int J Numer Meth Engng 48:1549–1570.
8. Ventura G (2006) Int J Numer Meth Engng 66:761–795.
9. Patzák B, Jirásek M (2003) Engrg Fracture Mech 70:957-977.
10. Areias P M A, Belytschko T (2006) Int J Numer Meth Engng 66:878–910.
11. Stazi F, Budyn E, Chessa J, Belytschko T (2003) Comput Mech 31:38–48.

OTHER DISCRETIZATION METHODS

Numerical determination of crack stress and deformation fields in gradient elastic solids

G.F. Karlis[1], S.V. Tsinopoulos[2], D. Polyzos[3] and D.E. Beskos[4]

[1] Department of Mechanical and Aeronautical Engineering, University of Patras, GR-26500 Patras, Greece gkarlis@mech.upatras.gr
[2] Department of Civil Engineering, University of Patras, GR-26500 Patras, Greece stsinop@upatras.gr
[3] Department of Mechanical and Aeronautical Engineering, University of Patras, GR-26500 Patras, Greece and Institute of Chemical Engineering and High Temperature Chemical Process, Patras, Greece polyzos@mech.upatras.gr
[4] Department of Civil Engineering, University of Patras, GR-26500 Patras, Greece d.e.beskos@upatras.gr

Summary. A boundary element method is developed for fracture analysis of gradient elastic 2-D solids under static loading. A simple version of Mindlin's general theory of gradient elastic materials is employed and the two required boundary integral equations, one for displacements and the other for its normal derivative are presented. Use is made of the fundamental solution of the problem and this leads to a formulation that requires only a boundary discretization. Two representative numerical examples are presented to illustrate the method, demonstrate its accuracy and efficiency and assess the gradient effect on the response. The first deals with a mode I crack, while the second with a mixed mode (I & II) crack. For the second case the proposed method is used in conjunction with the method of subregions. The method is employed with regular and regular plus special (near the crack tip) boundary elements. The gradient effect consists of modifying both the displacement and the stress field around the crack tip and resulting in a response which is more physically acceptable than the one coming from the classical theory of elasticity.

Key words: gradient elasticity, fracture mechanics, BEM, variable-order singularity element

1 Introduction

The mechanical behavior of linear elastic materials with microstructure, such as polymers, polycrystals or granular materials, cannot be adequately described by the classical elasticity theory and the concept of the local stress. When the microstructural effects are important, the state of stress is defined in

Alain Combescure et al. (eds.), IUTAM Symposium on Discretization Methods for Evolving Discontinuities, 365–380.

a non-local manner and the material behavior can be macroscopically modeled by using higher order strain gradient, micropolar or couple stress theories.

Among these theories, the simple gradient elastic theory with just one or two constants in addition to the two classical Lamè constants has received considerable attention during the last fifteen years or so. This theory is a special case of the simplified versions of the general gradient elasticity theory of Mindlin [1, 2]. These simple gradient elasticity theories have been successfully used for the analytic solution of various boundary value problems and have demonstrated the elimination of singularities or discontinuities of classical elasticity, their ability to capture size and edge effects and wave dispersion in cases where this was not possible in the classical elasticity framework (Aifantis [3], Ru and Aifantis [4], Vardoulakis and Sulem [5], Exadaktylos and Vardoulakis [6]).

In linear elastic fracture analysis one encounters large strain gradients near the crack tip and gradient elasticity theory appears to be the ideal theory for studying the strain and stress fields near the crack tip at the microscale. One can mention here the analytical works of Vardoulakis et al [7], Vardoulakis and Exadaktylos [8], Exadaktylos [9], Shi et al [10], Zhang et al [11], Fannjiang et al [12] and Georgiadis [13] on the determination of strain and stress fields around the tip of mode I, II and III cracks under conditions of plane strain or anti-plane strain. In all these works no computation of stress intensity factors (SIF) has been reported, because of the complexity of the problem. It is obvious that for complex gradient elastic fracture mechanics problems use of numerical methods of solution is imperative.

Shu et al [14] and Amanatidou and Aravas [15] have used the finite element method (FEM) for solving two-dimensional (2-D) elastostatic problems in the framework of the general theories of Mindlin. For the same purpose Tang et al [16] proposed a meshless local Petrov-Galerkin (MLPG) methodology, while Polyzos et al [17] and Tsepoura et al [18, 19] developed a boundary element method (BEM) for solving two and three dimensional elastostatic problems in the context of the simple gradient elastic theory.

In the present work, the gradient elastic BEM proposed by Polyzos et al [17] and Tsepoura et al [18] is employed for the solution of 2-D mode-I and mixed mode (I & II) fracture mechanics problems. This is accomplished by a displacement type of formulation in connection with the multiregion approach. For the other types of BEM formulations for fracture mechanics problems one can consult [20–22]. As it is explained in [7–13], near the tip of the crack displacements and strains are regular, while double stresses and total stresses exhibit a singular behavior of order $r^{-1/2}$ and $r^{-3/2}$, respectively, with r being the distance from the tip. Adopting the idea of variable-order singularity boundary elements around the tip of the crack for the evaluation of the corresponding stress intensity factor (SIF) [23], a new special variable-order singularity discontinuous element is proposed here for the treatment of singular fields around the tip of the crack. The paper is organized as follows: the constitutive equations, the classical as well as the non-classical boundary

conditions and the integral representation of the considered gradient elastic fracture mechanics problem are presented in the following section. The proposed new special variable-order singularity discontinuous element is illustrated next in the third section. The numerical implementation of the problem and the solution procedure of the utilized BEM code are explained in brief in the fourth section. Finally, the mode-I and the mixed mode (I & II) crack problems are solved and the corresponding crack profiles and SIFs are presented and discussed. The paper closes with a list of conclusions.

2 Simple gradient elasticity theory and integral formulation

In this section the equation of equilibrium, the corresponding boundary conditions and the integral representation of a gradient elastic problem are presented in brief. More details can be found in [18]. Mindlin [2], considering isotropic materials and a special case of his general theory, where the macroscopic strain coincides to micro-deformation, proposed a modification of Hookes law involving only five constants. A simpler and mathematically more tractable constitutive equation can be obtained by further reducing the number of constants to just three. In that case one has the constitutive equations

$$\begin{aligned} \tilde{\boldsymbol{\sigma}} &= \tilde{\boldsymbol{\tau}} + \tilde{\mathbf{s}} \\ \tilde{\boldsymbol{\tau}} &= 2\mu\tilde{\mathbf{e}} + \lambda\left(\nabla\cdot\bar{\mathbf{u}}\right)\tilde{\mathbf{I}} \\ \tilde{\mathbf{e}} &= \left(\nabla\bar{\mathbf{u}} + \bar{\mathbf{u}}\nabla\right)/2 \\ \tilde{\boldsymbol{\mu}} &= g^2\nabla\tilde{\boldsymbol{\tau}} \\ \tilde{\mathbf{s}} &= -\nabla\cdot\tilde{\boldsymbol{\mu}} = -g^2\nabla^2\tilde{\boldsymbol{\tau}} \end{aligned} \tag{1}$$

where ∇ is the gradient operator, ∇^2 is the Laplacian, $\tilde{\boldsymbol{\sigma}}$ is the total stress tensor, $\tilde{\boldsymbol{\tau}}$ and $\tilde{\mathbf{s}}$ are the so-called by Mindlin, Cauchy stress tensor and relative stress tensor, respectively, λ and μ are the Lamè constants, $\bar{\mathbf{u}}$ is the displacement vector, $\tilde{\mathbf{e}}$ is the strain tensor and g is the volumetric strain gradient energy coefficient, the only constant which relates the microstructure with the macrostructure and represents the characteristic length of the material.

Consider a finite 2-D gradient elastic body of volume V surrounded by a smooth boundary S. According to Mindlin's [1,2] theory in conjunction with the assumption of zero body forces acting on the body, the static equation of equilibrium reads

$$\nabla\cdot\left(\tilde{\boldsymbol{\tau}} - \nabla\cdot\tilde{\boldsymbol{\mu}}\right) = \mathbf{0} \tag{2}$$

and is accompanied by the classical boundary conditions

$$\begin{aligned} &\mathbf{u}\left(\mathbf{x}\right) = \mathbf{u_0},\ \mathbf{x}\in S_1 \text{ and} \\ &\mathbf{P}\left(\mathbf{x}\right) = \mathbf{P_0},\ \mathbf{x}\in S_1, \text{ with } S_1\cup S_2 = S \end{aligned} \tag{3}$$

and the non-classical ones

$$\mathbf{q}(\mathbf{x}) = \frac{\partial \mathbf{u}}{\partial n} = \mathbf{q_0},\ \mathbf{x} \in S_3 \text{ and}$$
$$\mathbf{R}(\mathbf{x}) = \hat{\mathbf{n}} \cdot \tilde{\boldsymbol{\mu}} \cdot \hat{\mathbf{n}} = \mathbf{R_0},\ \mathbf{x} \in S_4, \text{ with } S_3 \cup S_4 = S \tag{4}$$

where $\hat{\mathbf{n}}$ is the normal unit vector on S, $\tilde{\boldsymbol{\tau}}$ the classical elastic stress tensor, $\tilde{\boldsymbol{\mu}}$ the third order tensor of double forces per unit area, $\mathbf{P}$ the external surface tractions, $\mathbf{R}$ the surface double stress tractions and $\mathbf{P_0}$, $\mathbf{u_0}$, $\mathbf{R_0}$ and $\mathbf{q_0}$ prescribed values.

Adopting the above simplified theory of Mindlin and inserting the constitutive equations (1) into eq.(2) one obtains the following equation of equilibrium of a gradient elastic continuum in terms of the displacement field $\bar{\mathbf{u}}$:

$$\mu\nabla^2\bar{\mathbf{u}} + (\lambda+\mu)\nabla\nabla\cdot\bar{\mathbf{u}} - g^2\nabla^2\left(\mu\nabla^2\bar{\mathbf{u}} + (\lambda+\mu)\nabla\nabla\cdot\bar{\mathbf{u}}\right) = \mathbf{0} \tag{5}$$

The fundamental solution of the present gradient elastostatic problem is of the form [18]

$$\tilde{\mathbf{u}}^*(r,\mu,\nu,g) = \frac{1}{16\pi\mu(1-\nu)}\left[\Psi(r,\nu,g)\,\tilde{\mathbf{I}} - X(r,g)\,\hat{\mathbf{r}}\otimes\hat{\mathbf{r}}\right] \tag{6}$$

where ν is the Poisson ratio, $\hat{\mathbf{r}}$ the unit vector in the direction $\mathbf{r} = \mathbf{y} - \mathbf{x}$ and X, Ψ scalar functions given by the relations

$$X = -2 + \frac{8g^2}{r^2} - 4K_2\left(\frac{r}{g}\right) \tag{7}$$

$$\Psi = -2(3-4\nu)\ln r + \frac{4g^2}{r^2} - 2(3-4\nu)K_0\left(\frac{r}{g}\right) - 2K_2\left(\frac{r}{g}\right) \tag{8}$$

with $K_0(\cdot)$ and $K_2(\cdot)$ being the modified Bessel functions of the second kind and zero and second order, respectively.

For a smooth boundary S the integral representation of the problem has the form [18]

$$\tilde{\mathbf{c}}(\mathbf{x})\cdot\mathbf{u}(\mathbf{x}) + \int_S \left\{\tilde{\mathbf{p}}^*(\mathbf{x},\mathbf{y})\cdot\mathbf{u}(\mathbf{y}) - \tilde{\mathbf{u}}^*(\mathbf{x},\mathbf{y})\cdot\mathbf{p}(\mathbf{y})\right\} dS_y = \int_S \left\{\frac{\partial \mathbf{u}^*(\mathbf{x},\mathbf{y})}{\partial n_y}\cdot\mathbf{R}(\mathbf{y}) - \tilde{\mathbf{R}}^*(\mathbf{x},\mathbf{y})\cdot\mathbf{q}(\mathbf{y})\right\} dS_y \tag{9}$$

where $\tilde{\mathbf{u}}^*(\mathbf{x},\mathbf{y})$ is the fundamental solution given by eq.(6), $\tilde{\mathbf{p}}^*(\mathbf{x},\mathbf{y})$ and $\tilde{\mathbf{R}}^*(\mathbf{x},\mathbf{y})$ are the fundamental traction and double stress traction tensors, respectively and $\tilde{\mathbf{c}}(\mathbf{x})$ is the well-known jump tensor being equal to $1/2\tilde{\mathbf{I}}$ for $\mathbf{x} \in S$ and equal to $\tilde{\mathbf{I}}$ when $\mathbf{x} \in V \cap S$. Observing eq.(9), one realizes that this equation contains four unknown vector fields, $\mathbf{u}(\mathbf{x})$, $\mathbf{P}(\mathbf{x})$, $\mathbf{R}(\mathbf{x})$ and $\mathbf{q}(\mathbf{x}) = \partial\mathbf{u}/\partial n$. Thus, after the satisfaction of the classical and non-classical boundary conditions, the evaluation of the unknown fields requires

the existence of one more integral equation. This integral equation is obtained by applying the operator $\partial/\partial n$ on eq.(9) and has the form

$$\tilde{\mathbf{c}}\left(\mathbf{x}\right)\cdot\frac{\partial\mathbf{u}\left(\mathbf{x}\right)}{\partial n_x}+\int_S\left\{\frac{\partial\tilde{\mathbf{p}}^*\left(\mathbf{x},\mathbf{y}\right)}{\partial n_x}\cdot\mathbf{u}\left(\mathbf{y}\right)-\frac{\partial\tilde{\mathbf{u}}^*\left(\mathbf{x},\mathbf{y}\right)}{\partial n_x}\cdot\mathbf{p}\left(\mathbf{y}\right)\right\}dS_y=$$
$$\int_S\left\{\frac{\partial^2\tilde{\mathbf{u}}^*\left(\mathbf{x},\mathbf{y}\right)}{\partial n_x\partial n_y}\cdot\mathbf{R}\left(\mathbf{y}\right)-\frac{\partial\tilde{\mathbf{R}}^*\left(\mathbf{x},\mathbf{y}\right)}{\partial n_x}\cdot\mathbf{q}\left(\mathbf{y}\right)\right\}dS_y \quad (10)$$

All the kernels appearing in the integral equations (9) and (10) are given explicitly in [18].

Integral equations (9) and (10) accompanied by the classical and non-classical boundary conditions form the integral representation of the general gradient elastic boundary value problem.

3 Discontinuous line element of variable order singularity

According to [7–13], the fields $\mathbf{u}$, $\mathbf{q}$, $\mathbf{R}$ and $\mathbf{P}$ near the crack tip vary as $r^{3/2}$, $r^{1/2}$, $r^{-1/2}$ and $r^{-3/2}$ respectively, with r being the distance from the tip. As it is well known, the elements used in a classical BEM formulation interpolate the unknown fields either linearly or quadratically and therefore the behavior of the fields around the crack tip can never be represented correctly. In the present work, adopting the idea of using variable-order singularity boundary elements around the tip of the crack for the description of the near tip behavior and the evaluation of the corresponding SIFs [23], a new special variable-order of singularity discontinuous element is proposed. The advantage of this approach is that the fields around the tip of the crack are treated in a unified manner.

In this special element, the functional nodes are identical to those of a classical discontinuous three-noded quadratic line element, with the discontinuous side residing always at the crack tip. The main advantage of using discontinuous elements is that no functional nodes are located at the tip of the crack and thus, despite the singularity of $\mathbf{R}$ and $\mathbf{P}$ at the tip, the field nodal values are finite and can be easily computed.

As shown in 1, the tip of the crack can be located either at $\xi=-1$ or at $\xi=1$. In order to unify these two possible cases the linear transformation.

$$p=\frac{(1+c\xi)}{2} \quad (11)$$

is introduced with $c=\pm1$ for the tip residing at $\xi=\mp1$, respectively. Noting that $p\in[0,1]$, the tip of the crack is always located at $p=0$. The fields, in terms of the asymptotic solutions [23], can be expressed as

$$\mathbf{F}=\mathbf{K_1}r^{\lambda_1}+\mathbf{K_2}r^{\lambda_2}+C \quad (12)$$

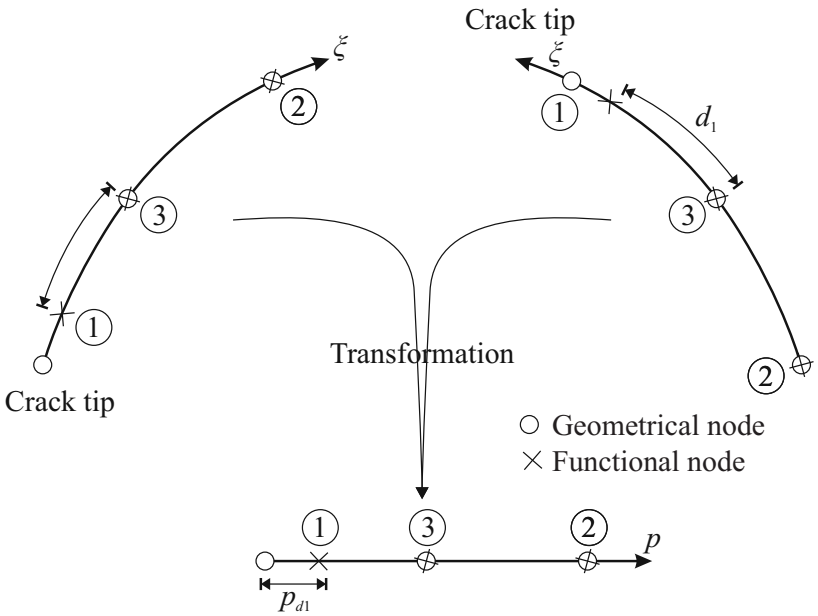


Fig. 1. Variable-order singularity discontinuous element and its transformation

Table 1. Orders of magnitude of the asymptotic fields

F	λ_1	λ_2
u	3/2	1
q	1/2	1
R	−1/2	1
P	−3/2	−1/2

with the symbol **F** representing **u**, **q**, **R** or **P** and λ_1, λ_2 taking values those in table 1. Writing eq.(12) for the three functional nodes of the variable-order of singularity discontinuous element one obtains the linear system of equations

$$\begin{aligned} r = p_{d_1} L: \ & \mathbf{K_1} \left(L p_{d_1}\right)^{\lambda_1} + \mathbf{K_2} \left(L p_{d_1}\right)^{\lambda_2} + \mathbf{C} = \mathbf{F_1} \\ r = \frac{L}{2}: \ & \mathbf{K_1} \left(\frac{L}{2}\right)^{\lambda_1} + \mathbf{K_2} \left(\frac{L}{2}\right)^{\lambda_2} + \mathbf{C} = \mathbf{F_3} \\ r = L: \ & \mathbf{K_1} \left(L\right)^{\lambda_1} + \mathbf{K_2} \left(L\right)^{\lambda_2} + \mathbf{C} = \mathbf{F_2} \end{aligned} \tag{13}$$

where $\mathbf{F_1}$, $\mathbf{F_2}$ and $\mathbf{F_3}$ are the nodal values of the field, p_{d_1} is the local coordinate of the discontinuous functional node and L is the length of the element. In the present work p_{d_1} is considered to be equal to 1/6. The solution of the linear system (13) yields the values for the parameters $\mathbf{K_1}$, $\mathbf{K_2}$ and $\mathbf{C}$ as functions of the nodal values $\mathbf{F_1}$, $\mathbf{F_2}$ and $\mathbf{F_3}$. Substituting $\mathbf{K_1}$, $\mathbf{K_2}$ and $\mathbf{C}$

into eq.(12), rearranging with respect to the nodal values $\mathbf{F_1}$, $\mathbf{F_2}$ and $\mathbf{F_3}$ and taking into account that the distance from the tip of the crack is $r = Lp$, the field $\mathbf{F}$ can be written as

$$\mathbf{F} = N_1\mathbf{F_1} + N_2\mathbf{F_2} + N_3\mathbf{F_3} \tag{14}$$

with the interpolation functions N_i being of the form

$$\begin{aligned}
N_1(p) &= \frac{2^{\lambda_1}\left[\left(1-2^{\lambda_2}\right)p^{\lambda_1}-1\right]-2^{\lambda_2}\left[\left(1-2^{\lambda_1}\right)p^{\lambda_2}-1\right]}{2^{\lambda_1}\left[\left(1-2^{\lambda_2}\right)p_{d_1}^{\lambda_1}-1\right]-2^{\lambda_2}\left[\left(1-2^{\lambda_1}\right)p_{d_1}^{\lambda_2}-1\right]} \\
N_2(p) &= \frac{-2^{\lambda_1}\left[\left(1-2^{\lambda_2}p_{d_1}^{\lambda_2}\right)p^{\lambda_1}-p_{d_1}^{\lambda_1}\right]+2^{\lambda_2}\left[\left(1-2^{\lambda_1}p_{d_1}^{\lambda_1}\right)p^{\lambda_2}-p_{d_1}^{\lambda_2}\right]}{2^{\lambda_1}\left[\left(1-2^{\lambda_2}\right)p_{d_1}^{\lambda_1}-1\right]-2^{\lambda_2}\left[\left(1-2^{\lambda_1}\right)p_{d_1}^{\lambda_2}-1\right]} \\
N_3(p) &= \frac{2^{\lambda_1+\lambda_2}\left[\left(1-p_{d_1}^{\lambda_2}\right)p^{\lambda_1}-\left(1-p+{d_1}^{\lambda_1}\right)p^{\lambda_2}-p_{d_1}^{\lambda_1}+p_{d_1}^{\lambda_2}\right]}{2^{\lambda_1}\left[\left(1-2^{\lambda_2}\right)p_{d_1}^{\lambda_1}-1\right]-2^{\lambda_2}\left[\left(1-2^{\lambda_1}\right)p_{d_1}^{\lambda_2}-1\right]}
\end{aligned} \tag{15}$$

It can be easily verified that $\sum N_i = 1$ and for $\lambda_1 = 2$ and $\lambda_2 = 1$ the functions (15) coincide with the classical discontinuous quadratic element interpolation functions. The interpolation functions (15) as functions of the local coordinate p are displayed in figs (2–5) for the fields $\mathbf{u}$, $\mathbf{q}$, $\mathbf{R}$ and $\mathbf{P}$, respectively. Observing figs (4) and (5), it is obvious that for the fields $\mathbf{R}$ and $\mathbf{P}$ the interpolation functions (15) become singular when approaching the tip of the crack ($p \to 0$).

4 BEM procedure

The goal of the boundary element methodology is to solve numerically the well-posed boundary value problem consisting of the system of the two boundary integral equations (9) and (10) with $\tilde{\mathbf{c}}(\mathbf{x}) = 1/2\tilde{\mathbf{I}}$ together with the boundary conditions (3) and (4). To this end, the smooth boundary S is discretized into three-noded quadratic line isoparametric boundary elements, except at either side of the crack tip, where special elements of the type described in the previous section are employed. The discretized boundary integral equations (9) and (10) are collocated at all nodal points and the linear system of algebraic equations

$$\begin{bmatrix} \frac{1}{2}\tilde{\mathbf{I}}+\tilde{\mathbf{H}} & \tilde{\mathbf{K}} \\ \tilde{\mathbf{S}} & \frac{1}{2}\tilde{\mathbf{I}}+\tilde{\mathbf{T}} \end{bmatrix} \cdot \begin{Bmatrix} \mathbf{u} \\ \mathbf{q} \end{Bmatrix} = \begin{bmatrix} \tilde{\mathbf{G}} & \tilde{\mathbf{L}} \\ \tilde{\mathbf{V}} & \tilde{\mathbf{W}} \end{bmatrix} \cdot \begin{Bmatrix} \mathbf{P} \\ \mathbf{R} \end{Bmatrix} \tag{16}$$

is obtained, where the influence matrices $\tilde{\mathbf{H}}$, $\tilde{\mathbf{K}}$, $\tilde{\mathbf{S}}$, $\tilde{\mathbf{T}}$, $\tilde{\mathbf{G}}$, $\tilde{\mathbf{L}}$, $\tilde{\mathbf{V}}$ and $\tilde{\mathbf{W}}$ contain line integrals over the boundary elements involving products of fundamental kernels times shape functions and Jacobians [18]. Applying the boundary

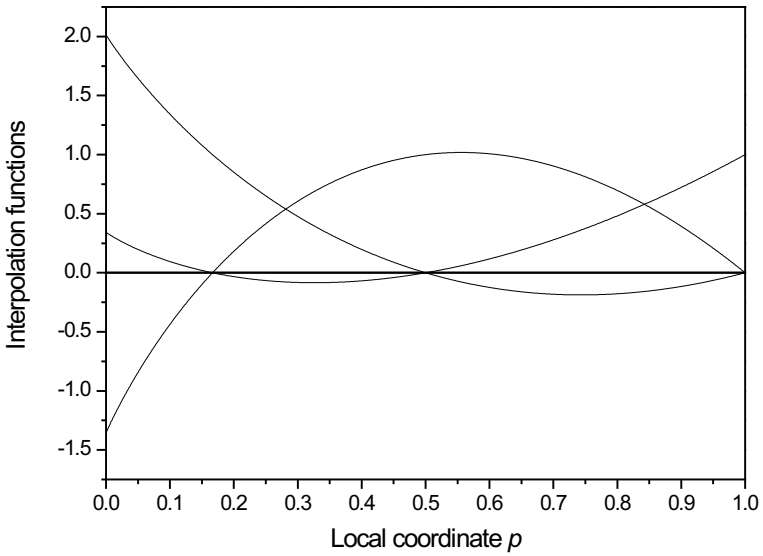


Fig. 2. Interpolation functions for displacements $\mathbf{u}$ ($\lambda_1 = 3/2$, $\lambda_2 = 1$)

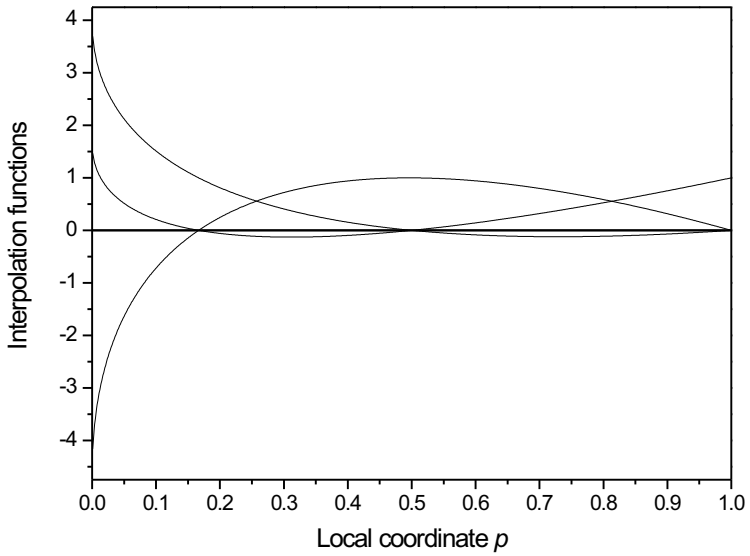


Fig. 3. Interpolation functions for displacement normal derivative $\mathbf{q}$ ($\lambda_1 = 1/2$, $\lambda_2 = 1$)

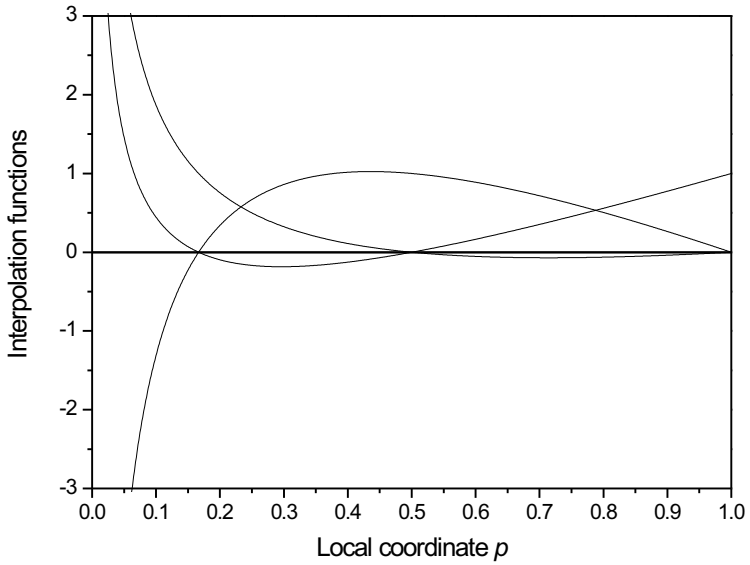


Fig. 4. Interpolation functions for double stresses $\mathbf{R}$ ($\lambda_1 = -1/2$, $\lambda_2 = 1$)

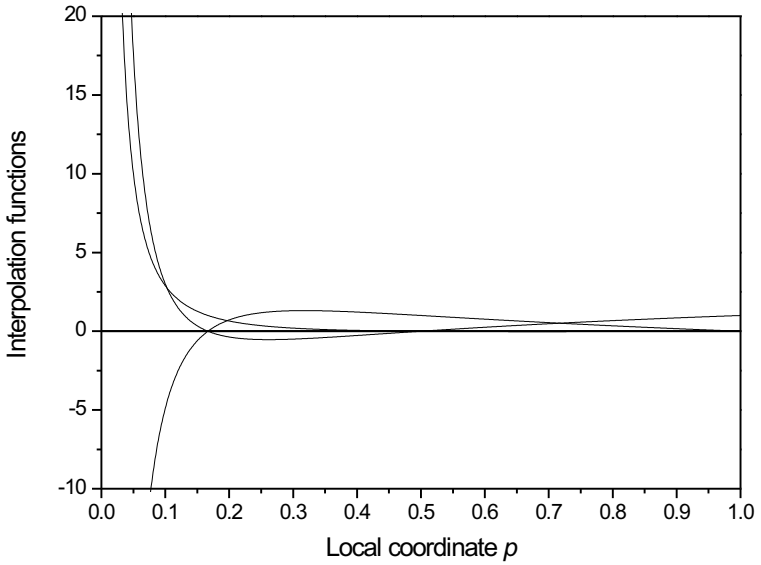


Fig. 5. Interpolation functions for tractions $\mathbf{P}$ ($\lambda_1 = -3/2$, $\lambda_2 = -1/2$)

conditions (3) and (4) and rearranging eq.(16), one produces the final linear system of algebraic equations of the form

$$\tilde{\mathbf{A}} \cdot \mathbf{X} = \mathbf{B} \tag{17}$$

where the vectors $\mathbf{X}$ and $\mathbf{B}$ contain all the unknown and known nodal components of the boundary fields, respectively.

The singular and hyper-singular integrals involved, are evaluated with high accuracy applying a methodology for direct treatment of Cauchy principal value and hyper-singular integrals [18], noting that an extra singularity due to the singular behavior of the interpolation functions (15) near the tip of the crack should be also taken into account. Finally, the linear system is solved via a typical LU-decomposition algorithm and the vector $\mathbf{X}$, comprising all the unknown nodal values of $\mathbf{u}$, $\mathbf{q}$, $\mathbf{R}$ and $\mathbf{P}$, is evaluated.

5 Computation of Stress Intensity Factors

Once the boundary value problem has been solved, the calculation of SIFs is being done via the three nodal traction values of the special elements.

Approaching the crack tip ($r \to 0$), the traction $\mathbf{P}$, according to eq.(12), admits a representation of the form

$$\mathbf{P} = \frac{\mathbf{K_1}\left(\mathbf{P_1}, \mathbf{P_2}, \mathbf{P_3}\right)}{\sqrt{2\pi}} \lim_{r\to 0} r^{-\frac{3}{2}} + \frac{\mathbf{K_2}\left(\mathbf{P_1}, \mathbf{P_2}, \mathbf{P_3}\right)}{\sqrt{2\pi}} \lim_{r\to 0} r^{-\frac{1}{2}} \tag{18}$$

where the components of the vectors $\mathbf{K_1}$ and $\mathbf{K_2}$ stand for the stress intensity factors corresponding to x and y directions according to the following relations:

$$\begin{aligned}
\mathbf{K_1} &= \left\{ \begin{array}{c} K_{1x} \\ K_{1y} \end{array} \right\} = \\
&= L^{\frac{3}{2}} \sqrt{2\pi} p_{d_1} \frac{\sqrt{p_{d_1}}\left(2-\sqrt{2}\right)\mathbf{P_1} + \left(\sqrt{2} - 2\sqrt{p_{d_1}}\right)\mathbf{P_2} - \left(\sqrt{2} - \sqrt{2p_{d_1}}\right)\mathbf{P_3}}{2 - \sqrt{2} + \left(-4+\sqrt{2}\right) p_{d_1} + 2p_{d_1}^{3/2}} \\
\mathbf{K_2} &= \left\{ \begin{array}{c} K_{2x} \\ K_{2y} \end{array} \right\} = \\
&= \sqrt{L}\sqrt{2\pi} \frac{p_{d_1}^{3/2}\left(\sqrt{2}-4\right)\mathbf{P_1} + \left(4p_{d_1}^{3/2} - \sqrt{2}\right)\mathbf{P_2} + \sqrt{2}\left(1 - p_{d_1}^{3/2}\right)\mathbf{P_3}}{2 - \sqrt{2} + \left(-4+\sqrt{2}\right) p_{d_1} + 2p_{d_1}^{3/2}}
\end{aligned} \tag{19}$$

6 Numerical examples

Two fracture problems, the first of mode-I and the second of mixed mode (I & II), are solved and the obtained crack profiles as well as the corresponding SIFs are presented and compared against those of classical elasticity.

Mode-I crack problem

A square gradient elastic plate with very small radius of curvature round corners (in order to have a smooth boundary) containing a central horizontal line crack and subjected to in-plane forces is analyzed. A uniform tensile traction $P_0 = 100\text{MPa}$, is applied normal to its top and bottom sides, as shown in fig.6. The crack length is chosen to be equal to $2a = 1\text{m}$ and the side of the square plate is $L = 16a$. The Young modulus and the Poisson ratio of the gradient elastic plate are $E = 210\text{GPa}$ and $= 0.2$, respectively. Due to the double symmetry of the problem, only one quarter of the plate is discretized, with the following boundary conditions along the axes of symmetry: $\mathbf{P}(0, y) = \mathbf{0}$ and $\mathbf{R}(0, y) = \mathbf{0}$ for $0 \leq y < a$, $u_y(0, y) = 0$ and $\mathbf{R}(0, y) = \mathbf{0}$ for $a \leq y \leq L/2$ and $u_x(x, 0) = 0$ and $\mathbf{R}(x, 0) = \mathbf{0}$ for $0 \leq x \leq L/2$.

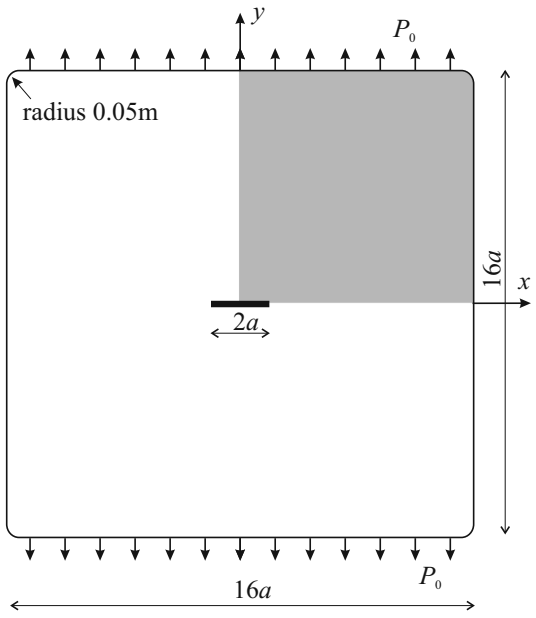


Fig. 6. Gradient elastic plate with a central diagonal line crack

Fig. 7 displays the upper-right-quarter of the crack opening displacement profile obtained by the present BEM for four different values of the material volumetric strain energy coefficient g (0.01, 0.1, 0.3, 0.5). In the same figure, the crack profile provided by the classical elasticity theory ($g = 0$) is also shown. The main conclusion here is that the crack profile of the gradient elastic case remains sharp at the crack tip and is not blunted as in the classical case. This cusp type of profile is identical to the one coming out of Barenblatts [24] cohesive zone theory. Also, it should be noticed that as the volumetric strain energy gradient coefficient g increases, the crack becomes stiffer.

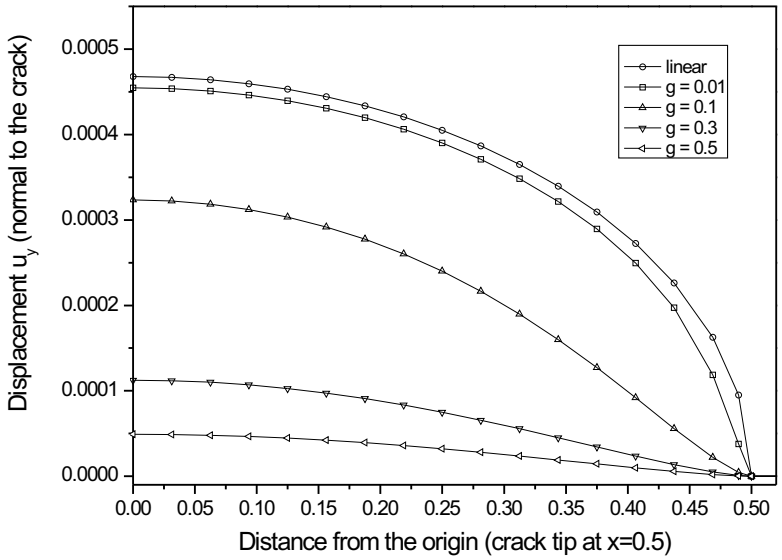


Fig. 7. Shape of mode-I crack for different values of the gradient coefficient g

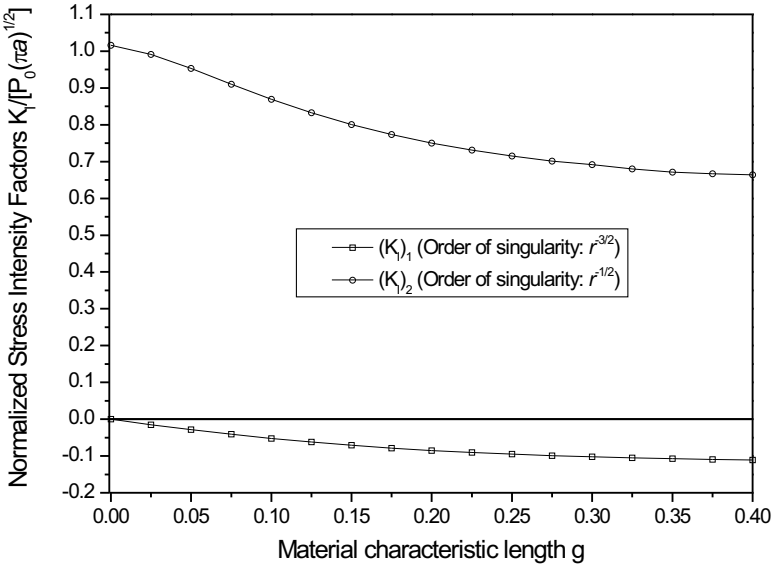


Fig. 8. SIFs with respect to g

In fig. 8 the two SIFs for a gradient elastic problem, $(K_I)_1$ and $(K_I)_2$, are plotted with respect to g. The interesting remark here is that the SIF $(K_I)_1$ tends to zero as the gradient coefficient g tends to zero. As a result of that, eq.(18) becomes $P_y = (K_I)_2 / \sqrt{2\pi} \lim_{r \to 0} r^{-1/2}$ with $(K_I)_2$ being the mode-I SIF as defined in classical elasticity theory.

Furthermore, the most important observation here is that the SIF $(K_I)_1$ takes only negative values. This means that in gradient elasticity the stresses near the crack tip not only go to infinity with a different order ($r^{-3/2}$) compared to the linear elasticity ($r^{-1/2}$), but are also compressive and not tensile as in classical elasticity. This explains the different shapes of the crack profile in gradient and classical elasticity theories, as shown in fig.7.

6.1 Mixed mode I & II crack problem

The square plate of the previous example with an inclined at an angle of 45^o central slant crack is analyzed here again by the proposed method. The problem domain is divided into two subregions, as shown in fig.9, which are both treated by the BEM and combined together through equilibrium and compatibility at their interface. This is necessary in view of the displacement based BEM employed here [20–22]. A traction based BEM would be also possible, but this is associated with higher order kernel singularities [20–22].

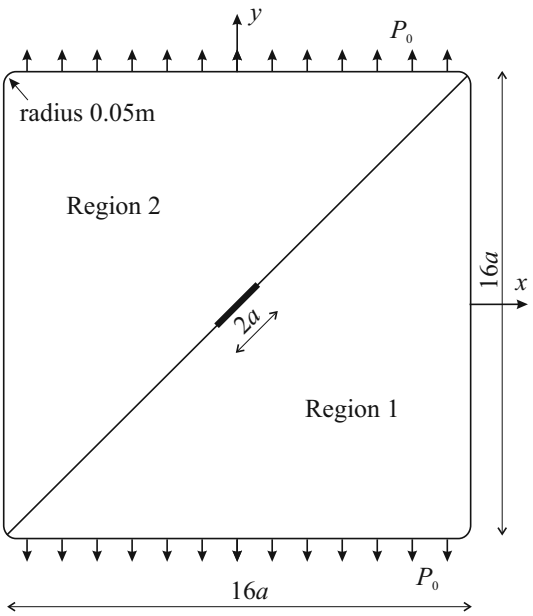


Fig. 9. Gradient elastic plate with a central diagonal line crack

In figs 10 and 11 the SIFs, $(K_I)_1$, $(K_{II})_1$ and $(K_I)_2$, $(K_{II})_2$, respectively are plotted as functions of g. Similarly to the mode-I case, the SIFs $(K_I)_1$, $(K_{II})_1$ tend to zero as the gradient coefficient g tends to zero and take only negative values. The crack opening displacement profile (not shown here) has the same shape as the one of the pervious mode-I case.

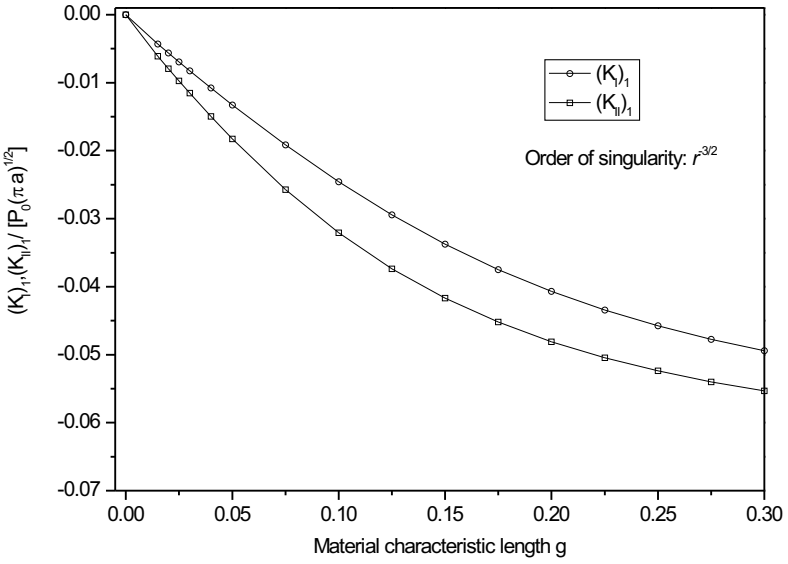


Fig. 10. SIFs, $(K_I)_1$, $(K_{II})_1$ with respect to g

7 Conclusions

On the basis of the preceding developments, the following conclusions can be drawn:

1. A displacement based BEM was employed for the solution of 2-D crack analysis problems involving gradient elastic material behavior. This approach requires, in general, the use of subregions but is associated with lower order kernel singularities than other approaches.
2. A new three-noded discontinuous line boundary element of variable singularity for the crack modeling has been developed. This special element can lead to the determination of the stress intensity factor(s) directly via its three nodal traction values after the determination of boundary tractions and displacements.

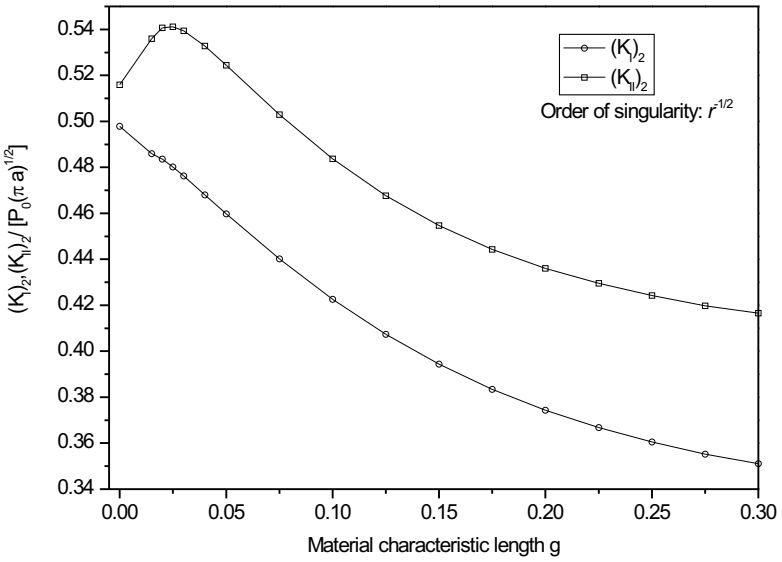


Fig. 11. SIFs, $(K_I)_2$, $(K_{II})_2$ with respect to g

3. The employed BEM leads to highly accurate results with a discretization which is restricted only to the boundaries and possible interfaces of the domains considered.
4. Stress and displacement fields around the crack tip in a gradient elastic plate are more physically acceptable than those of the case of classical elasticity.

8 Acknowledgements

We thank the European Social Fund (ESF), Operational Program for Educational and Vocational Training II (EPEAEK II), and particularly the Program PYTHAGORAS II, for funding the above work.

References

1. Mindlin RD (1964) Arch. Rat. Mech. Anal. 10:51–78
2. Mindlin RD (1965) Int. J. Solids Struct. 1:417–438
3. Aifantis EC (1992) Int. J. Engng. Sci. 30:1279–1299
4. Ru CQ, Aifantis EC (1993) Acta Mechanica 101:59–68
5. Vardoulakis I, Sulem J (1995) Bifurcation Analysis in Geomechanics. Blackie A & P - Chapman and Hall, London

6. Exadaktylos G, Vardoulakis I (2001) Technophysics 335:81–109
7. Vardoulakis I, Exadaktylos G, Aifantis E (1996) Int. J. Solids Struct. 33:4531–4559
8. Vardoulakis I, Exadaktylos G (1997) The asymptotic solution of anisotropic gradient elasticity with surface energy for a mode-II crack. In: IUTAM Symposium on non-linear singularities in deformation and flow, Haifa, Israel. Durban D (editor), luwer Academic Publishers, Dordrecht
9. Exadaktylos G (1998) Int. J. Solids Structures 35:421–456
10. Shi MX, Huang Y, Hwang KC (2000) J. Mech. Phys. Solids 48:2513–2538
11. Zhang L, Huang Y, Chen JY, Hwang KC (1998) Int. J. Fracture 92:325–348
12. Fannjiang AC, Chan YS, Paulino GH (2002) SIAM J. Appl. Math. 62:1066–1091
13. Georgiadis G (2003) J. Appl. Mech. ASME 70:517–530
14. Shu JY, King WE, Fleck NA (1999) Int. J. Num. Meth. Engng 44:373–391
15. Amanatidou E, Aravas N. (2002) Comp. Meth. Appl. Mech. Engng 191:1723–1751
16. Tang Z, Shen S, Atluri SN (2003) CMES: Comp. Modeling Engng Sci. 4:177–196
17. Polyzos D, Tsepoura KG, Tsinopoulos SV, Beskos DE. (2003) Comp. Meth. Appl. Mech. Engng 192:2845–2873
18. Tsepoura KG, Tsinopoulos SV, Polyzos D, Beskos DE (2003) Comp. Meth. Appl. Mech. Engng 192:2875–2907
19. Tsepoura KG, Polyzos D (2003) Comput. Mech. 32:89–103
20. Aliabadi MH. (1997) Appl. Mech. Rev. ASME 50:83–96
21. Beskos DE. (1997) Appl. Mech. Rev. ASME 50:149–197
22. Dominguez J, Ariza MP (2003) In: Boundary Element Advances in Solid Mechanics, Beskos DE, Maier G(eds), Springer, Wien
23. Lim KM, Lee KH, Tay AAO, Zhou W (2002) Int. J. Numer. Meth. Engng 55:293–316
24. Barenblatt GI. (1962) Adv. Appl. Mech. 7:55–129

The variational formulation of brittle fracture: numerical implementation and extensions

B. Bourdin

Department of Mathematics, Louisiana State University, Baton Rouge, LA 70803, USA (bourdin@math.lsu.edu)

Summary. This paper presents the implementation of a variational formulation of brittle fracture mechanics proposed by G.A. Francfort and J.-J. Marigo in 1998. The essence of the model relies on successive global minimizations of an energy with respect to any crack set and any kinematically admissible displacement field. We briefly present the model itself, and its variational approximation in the sense of Gamma–convergence. We propose a globally convergent and monotonically decreasing numerical algorithm. We introduce a backtracking algorithm whose solution satisfy a global optimality criterion with respect to the time evolution. We illustrate this algorithm with three dimensional numerical experiments. Then we present an extension of the model to crack propagation under thermal load and its numerical application to the quenching of glass.

Key words: Variational Model, Brittle Fracture, Crack Path Identification, Free Discontinuity Model, Gamma Convergence

Introduction

The work presented here is based an original approach proposed by G. Francfort and J.-J. Marigo in [18, 19]. While this model is still limited to quasi-static problems under fixed displacement boundary conditions, we extend it to account for body forces under some restrictions. The main virtue of the model we use is to remain largely compatible with Griffth theory, departing as little as possible to allow crack initiation or branching, path identification, and interactions between multiple cracks. However, these benefits have a cost in terms of complexity of the numerical implementation. The Francfort-Marigo formulation involves a *global* minimization of a total energy with respect to *any* admissible crack set and displacement field, and requires specialized numerical tools which we present in this article.

Alain Combescure et al. (eds.), IUTAM Symposium on Discretization Methods for Evolving Discontinuities, 381–393.

1 Francfort and Marigo's model for quasi-static brittle fracture

The brittle fracture formulation studied in this paper was first introduced by G. Francfort and J.J. Marigo, see [18, 17, 16, 12, 13] for a comprehensive presentation of the model. We will only briefly recall the essential points of this approach. As the emphasis of this paper is on the numerical implementation, we limit ourselves to a time discrete formulation. The reader will refer to [17, 16] for the time-continuous limit of the model. We differ from the original in that we consider body forces (under some restrictions) while the original formulation is restricted to fixed displacement boundary conditions.

In all that follows, we consider a two or three dimensional body with reference configuration $\Omega \subset \mathbb{R}^N$ $(N = 1, 2, 3)$ open, bounded with Lipschitz boundary. We assume that Ω can be partitioned into two connected subdomains Ω_F and Ω_D, corresponding respectively to a *fragile* and *ductile* materials with Hooke's law $\mathrm{A_F}$ and $\mathrm{A_D}$. We suppose that a part $\Omega_{D,0} \subset \Omega_D$ with non-null measure is clamped while $\partial\Omega \setminus \partial\Omega_{D,0}$ remains traction free.

The first key ingredient of the Francfort-Marigo model is to identify the cracks of the brittle material with the discontinuities of the displacement field. For that matter, we consider displacement fields in the space of Special Functions of Bounded Deformations (SBD). A detailed presentation of the space SBD is far beyond the scope of this article, the key point here is that any function $u \in SBD$ may be discontinuous, and that one can define its discontinuity set (or jump set) J_u.

We consider $P + 1$ time steps $0 \leq t^{(0)} < ... < t^{(P)} = T$. At each time step $t^{(p)}$, we apply a force $f(t^{(p)}; x)$ on Ω_D. In the sequel, we will assume that f depends linearly on t, that is that $f(t; x) := t.f(x)$.

The set of kinematically admissible displacement fields is

$$\mathcal{K}_A^{(p)} := \left\{ u \in SBD(\Omega)\,;\, u(x) = 0 \text{ a.e. in } \Omega_{D,0}\,;\, J_u \subset \bar{\Omega}_F\,;\, \|u\|_\infty \leq M \right\}, \quad (1)$$

where M is an arbitrarily large constant whose role is purely technical. Again we will refer the reader to [18, 17] for a complete exposition of the model.

In order to account for the *irreversible* nature of the fracture process, we define the *total* jump set at time step p of a sequence $\left(u^{(0)}, \ldots, u^{(P)}\right)$ by

$$\Gamma_u^{(p)} := \bigcup_{0 \leq s \leq p} J_{u^{(s)}}. \quad (2)$$

Note that we have trivially that $\Gamma_u^{(s)} \subset \Gamma_u^{(p)}$ for any $s \leq p$ so that the total crack set indeed grows monotonically.

The second key of Francfort Marigo is the *global minimization* of a total energy with respect to *any* admissible displacement field. For that, we define the bulk, surface, and total energies by

$$E^b(u(t^{(p)})) := \frac{1}{2}\int_{\Omega_F} \mathrm{A_F}\mathbf{e}(u) : \mathbf{e}(u)\, dx + \frac{1}{2}\int_{\Omega_D} \mathrm{A_D}\mathbf{e}(u) : \mathbf{e}(u) \; - 2f \cdot u\, dx, \tag{3}$$

$$E^s(u(t^{(p)})) := \gamma \mathcal{H}^{N-1}(\Gamma_u^{(p)}), \tag{4}$$

$$E(u(t^{(p)})) := E^b(u(t^{(p)})) + E^s(u(t^{(p)})), \tag{5}$$

where $\mathcal{H}^{N-1}$ denotes the $N-1$–dimensional Hausdorff measure (*i.e.* the length in two dimensions and surface in the three dimensional case), $\mathbf{e}(u)$ denotes the symmetrized gradient of u, and γ is the fracture toughness of the brittle material considered. At time $t^{(p)}$, the displacement field $u(t^{(p)})$ is solution of the *global minimization problem*:

$$\inf_{u \in \mathcal{K}_A^{(p)}} E(u), \tag{6}$$

and will denote by $E(t^{(p)})$ the total energy $E(u(t^{(p)}))$. Similarly, by $E^s(t^{(p)})$ and $E^b(t^{(p)})$, we refer to the surface and bulk energy of the solution of (6).

As the crack set at time t is given through a global minimization process among all possible crack states, the Francfort-Marigo model does not require *a priori* knowledge of the crack path. It does not even require the existence of an initial crack. It also does not assume smooth propagation of cracks (*i.e.* that the surface energy $E_s(t)$ is a continuous function of t). Indeed, as we will see in the numerical experiments, it is often the case that the total crack length is a discontinuous function of the time, a phenomenon we will refer to as *brutal* crack propagation.

2 Numerical implementation

In order to discretize the Francfort-Marigo functional, one needs to be able to approximate *any* function in SBD. This is by nature more complicated that building a discrete space allowing jumps across a known curve or surface. For that reason, the extended finite elements method is not easily applicable. This model also requires the ability to accurately approximate the location of the cracks, as well as their *length*, which may not be possible if the cracks are restricted to propagate along edges of faces in between elements like in a discontinuous Galerkin or cohesive finite element methods, for instance. Lastly, in light of [14], it is expected that in absence of singularity in the deformation field, cracks initiation will always be brutal. In particular, this means that sensitivity with respect to "small" cracks will never provide a descent direction toward a global minimizer of the Francfort-Marigo energy, in the case of "brutal" crack evolution.

Several methods have been proposed, based on discontinuous (see [21]) or adaptive (see [9, 24]) finite elements. Our previous experience with adaptive

finite elements is that the mesh adaption step can introduce artificial local minimizers and render the global minimization of E_ε practically impossible. The method which we present here relies on approximating the Francfort-Marigo energy, in the sense of Γ–convergence, by means of elliptic functionals. It is similar to the one proposed in [3, 4] for the Mumford-Shah functional, inspired by a now classical example in phase transition by [22, 23, 2] and extended in [6, 12, 13, 20].

2.1 Approximation of the Francfort-Marigo energy

Following [3, 4], we introduce a secondary variable $v \in W^{1,2}(\Omega)$, representing the crack in some sense, and for any $\varepsilon > 0$, $\eta_\varepsilon = o(\varepsilon)$, and $\alpha_\varepsilon = o(\varepsilon)$, we define

$$E_\varepsilon(u,v) := E^b_\varepsilon(u,v) + E^s_\varepsilon(v) := \frac{1}{2}\int_{\Omega_F} \left(v^2+\eta_\varepsilon\right) \mathrm{A_F}\mathbf{e}(u) : \mathbf{e}(u)\, dx + \frac{1}{2}\int_{\Omega_D} \mathrm{A_D}\mathbf{e}(u) : \mathbf{e}(u) \; - 2f\cdot u\, dx + \gamma \int_\Omega \frac{(1-v)^2}{4\varepsilon} + \varepsilon|\nabla v|^2\, dx. \quad (7)$$

We do not attempt to prove the Γ–convergence of E_ε to E here. However, it is a known result in the original setting of the Francfort-Marigo model (*i.e.* when considering fixed displacement boundary conditions) and does not seem to be difficult to adapt to this case (see [10, 12, 13, 20]). Using a classical compactness argument (see for instance [15, 11]), one obtains that the global minimizers of $(u^\varepsilon, v^\varepsilon)$ of E_ε converge to that of E, and that in some weak sense, the set $\{x \in \Omega\,;\, v^\varepsilon(x) \le \alpha_\varepsilon\}$ converges to J_u.

In order to account for the irreversibility, we define

$$K^{(p)}_\varepsilon := \left\{ x \in \bar{\Omega}\ ;\ v^{(p)} \le \alpha_\varepsilon \right\}. \quad (8)$$

At each time step, we seek for $\left(u^{(p)}_\varepsilon, v^{(p)}_\varepsilon\right)$ solution of the problem

$$\begin{cases} \inf & E_\varepsilon(u,v). \\ u \in \mathcal{K}^{(p)}_A & \\ v = 0 \text{ on } K^{(p)}_\varepsilon. & \end{cases} \quad (9)$$

The last step in view of the numerical implementation of Francfort and Marigo's brittle fracture model is its discretization. For the Mumford-Shah problem, one can consider a discretized version $F_{\varepsilon,h}$ of F_ε by means of linear finite elements. Provided that the mesh size h is such that $h \ll \varepsilon$, it is known that $F_{\varepsilon,h}$ Γ–converges to F (see [5, 7]). Extending this result to our case does not seem to present any difficulty but is again not the scope of this study. We will take for granted that the restriction $F_{\varepsilon,h}$ of F_ε to discrete functions on a linear finite element space Γ–converges to F. This relation between the mesh

size and the regularization parameter can be an issue. However, a careful study of the Γ–convergence results in [5, 7] reveal that the relation $h \ll \varepsilon$ needs only to be verified "close" to the cracks. Unfortunately, finding the location of the cracks is the essence of the problem! In the numerical results we present, we used uniformly fine meshes, which can result very large meshes (typically, 10,000 to 50,000 elements in two dimensions and 200,000 to 1,500,000 elements in three dimensions).

2.2 Minimization methods

A major hurdle in the way of the minimization of (7) is its non-convexity. However, it is easy to see that E_ε is convex with respect to each of its arguments *separately*, and can therefore be iteratively minimized with respect to u and v. In the experiments presented later we used the following alternate minimization algorithm, δ being a fixed tolerance parameter:

Algorithm 1 *The alternate minimization algorithm*

1: *Let* $i = 0$ *and* $v_0 := v_\varepsilon^{(p-1)}$ *if* $p > 0$ *or* $v_0 = 1$ *if* $p = 0$.
2: ***repeat***
3: $\quad i \leftarrow i + 1$
4: $\quad$ *Compute* $u_i := \operatorname{argmin}_u F_\varepsilon(u, v_{i-1})$ *with* $u_i = 0$ *on* $\Omega_{D,0}$.
5: $\quad$ *Compute* $v_i := \operatorname{argmin}_v F_\varepsilon(u_i, v)$ *under the constraint* $v_i = 0$ *on* $K_\varepsilon^{(p-1)}$
6: ***until*** $\|v_i - v_{i-1}\|_\infty \leq \delta$
7: *Set* $u_\varepsilon^{(p)} := u_i$ *and* $v_\varepsilon^{(p)} := v_i$

It can be proved (see [8]) that in cases where the crack propagation is smooth, this algorithm converges to the global minimizer, provided that the time discretization is fine enough. In cases where cracks propagate brutally, this algorithm can only be proved to converge to a critical point of E_ε, which may be a local (or a global) minimizer, but also a saddle point for E_ε. In some cases, the local minimizers of E_ε can be proved to converge to local minimizers of E, but no such a thing can be said of its saddle points.

A full treatment of saddle points would require studying the stability of critical points, and has not been implemented yet. In order to detect some saddle points and local minimizers, we implemented a backtracking in time algorithm, based on necessary condition for optimality of the crack evolution with respect to time. In the sequel, we assume that the boundary condition grows linearly with time, *i.e.* that $g(t; x) := tg(x)$. If this is the case, it is easy to see that if $(u^{(p)}, v^{(p)})$ is admissible for E_ε at time step $t^{(p)}$, then for any $0 \leq r \leq p$, $\left(\frac{t^{(r)}}{t^{(p)}} u^{(p)}, v^{(p)}\right)$ is admissible for $t^{(r)}$. Noticing then that $E_\varepsilon^b\left(\frac{t^{(r)}}{t^{(p)}} u^{(p)}, v^{(p)}\right) = \left(\frac{t^{(r)}}{t^{(p)}}\right)^2 E_\varepsilon^b(u^{(p)}, v^{(p)})$, we see that if $u^{(r)}$ and $u^{(p)}$ are global minimizers of E at time steps $t^{(r)}$ and $t^{(r)}$, then one necessarily has that

$$E_\varepsilon(u^{(r)}, v^{(r)}) \leq \left(\frac{t^{(r)}}{t^{(p)}}\right)^2 E^b_\varepsilon(u^{(p)}, v^{(p)}) + E^s_\varepsilon(v^{(p)}), \text{ for any } 0 \leq r \leq p. \quad (10)$$

In order to ensure that condition (10) is satisfied at each time step, we implemented the following backtracking algorithm, δ_ε being a small tolerance parameter:

Algorithm 2 *The backtracking algorithm*

1: $v_0 \leftarrow 1$
2: ***repeat***
3: *Compute* $(u^{(p)}, v^{(p)})$ *using the alternate minimization algorithm initialized with* v_0.
4: *Compute the bulk energy* $E^b_\varepsilon(u^{(p)}, v^{(p)})$ *and the surface energy* $E^s_\varepsilon(u^{(p)}, v^{(p)})$
5: ***for*** $r = 1$ *to* $p - 1$ ***do***
6: ***if*** $E_\varepsilon(u^{(r)}, v^{(r)}) - \left(\frac{t^{(r)}}{t^{(p)}}\right)^2 E^b_\varepsilon(u^{(p)}, v^{(p)}) - E^s_\varepsilon(v^{(p)}) \geq \delta_\varepsilon$ ***then***
7: $v_0 \leftarrow v^{(p)}$
8: $p \leftarrow r$
9: ***return*** *to 3:*
10: ***end if***
11: ***end for***
12: $v_0 \leftarrow v^{(p)}$
13: $p \leftarrow p + 1$
14: ***until*** $p = P$

3 Numerical Experiments

We present some numerical experiments based on the formulation and algorithms presented above. We consider a three dimensional cylinder of radius 2 and length 10, along the z–axis reinforced in its center by a ductile shaft of radius .5 (depicted in light grey in Figure 1) capped at each end by a rigid reinforcement (in black in Figure 1). The elastic moduli of both materials are $E = 1$ and $\nu = .2$, while the fracture toughness is normalized to $\gamma = 1$. The bottom cap is clamped while on the top one, a body force is applied along the positive z direction. On the brittle material, we initiated a hairline crack along a disk of radius .4 centered at the edge of the outer cylinder, and forming an angle of 30 degrees with the xy–plane.

The domain is meshed using 130,950 nodes and 584,150 linear tetrahedral elements. The magnitude of the force varies linearly between .5 and 5.5 in 250 time steps.

Figure 2 represent the evolution of the total and surface energies between the times $t = 1.4$ and $t = 1.9$. The location of the crack set at various times is shown in Figure 3. In these figures, the part of the domain corresponding to the britle material has been removed, and the ductile reinforcement is transparent. The upper and lower rows of figures correspond to the same

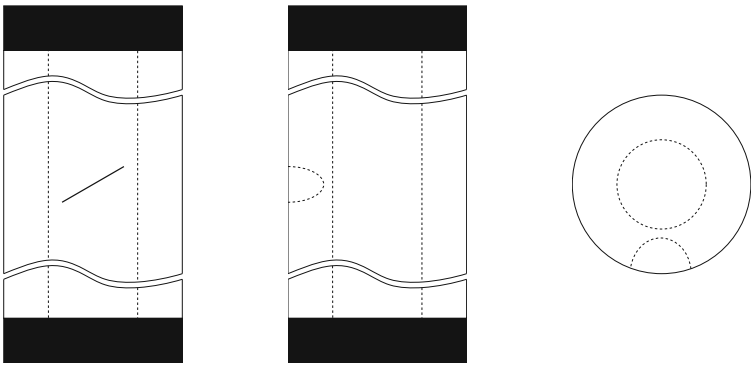


Fig. 1. Computational domain. Front, side and top views

time steps seen from different angles. The crack is represented by plotting the iso-volume $v \leq 5.0\text{E} - 2$.

The outcome of the alternate minimizations, combined with the backtracking is as follows:

- For $0 \leq t \leq 1.76$, no crack appears, however, the numerical solution exhibits some spurious surface energy (the lower branch of the dotted curve on Figure 2–right).
- At $t = 1.7$, the minimization algorithm bifurcates towards a cracked solution, and the surface energy jumps from 2.48 to 6.49. However, the total energy jumps from -12.5 to -14.2, which is forbidden by (10). The algorithm backtracks then to $t = 1.5$, time of the first violation of (10).
- Restarting from $t = 1.5$, alternate minimization lead to a smooth propagation of the crack until $t = 1.82$.
- At $t = 1.84$, the minimization algorithm bifurcates again towards a solution with an helicoidal crack. Again, the total energy jumps, which is in contradiction with (10), and the algorithm backtracks to $t = 1.64$
- At $t \geq 1.64$, the crack continue to grow, this time along the interface separating the fiber and the matrix (two rightmost figures in Figure 3). The evolution is smooth and the surface energy is continuous.

The final evolution can be summarized as follows:

- For $0 \leq t \leq 1.50$, no crack appears.
- At $t = 1.5$, a crack appears brutally starting from the existing notch (see Figure 3–left). The surface energy jumps (see the first jump of the plain curve in Figure 2–right).
- This crack propagates slowly for $1.50 \leq t \leq 1.62$ (see the slow growth of the plain curve in Figure 2–right) until it reaches the configuration in Figure 3 (second from the left).
- At $t = 1.64$, the crack propagates brutally again along a helicoidal path (Figure 3 center). The surface energy jumps again (second jump of the

plain curve in Figure 2–right). The crack spreads along the entire width of the brittle part, from the reinforcement until the edge of the matrix.

- For $t > 1.64$, a second branch of the crack develops along the reinforcement–matrix interface.

The total final and surface energy are represented by the plain lines in Figure 2. Using the backtracking algorithm, the total energy is continuous, (which is mandated by (10)), while solutions obtained without backtracking can enjoy discontinuities (see [24] Figure 11 or [10] Section 3.2). While we can still not guaranty that our solution corresponds to a global minimizer of the Francfort-Marigo energy, its energty is certainly less that that of a solution obtained without backtracking step (which would correspond to the upper envelope of the dashed curve in Figure 2–left).

This experiment also illustrates another strength of the Francfort-Marigo model and of our implementation. Recall again that in this experiment, the crack path (which is far from obvious) was not not known *a priori*. Using the Francfort-Marigo model, we were able to compute the crack path, as well as the position of the crack along this path. By using the approximation in the sense of Γ–convergence, we were able to represent complicated geometries. Note how the branch growing along the interface and the helix crack merge at $t = 3.08$. One of the virtue of our representation of the crack set is that it allows for such complicated change of topology without any special numerical treatment.

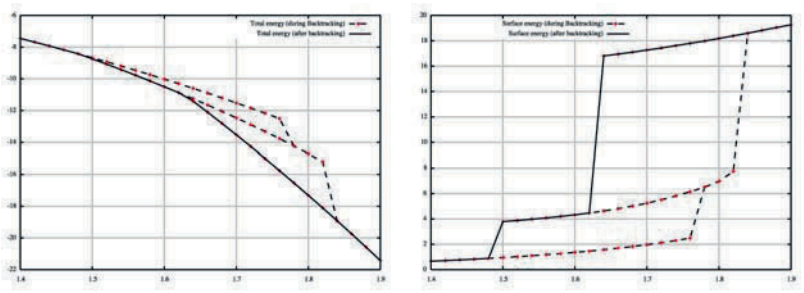


Fig. 2. Total and surface energy as a function of the load

4 Extensions to thermal loads

We present here some preliminary results on the extension of the Francfort-Marigo model to thermal loads.

In all that follows, we consider a *given* temperature field $\theta(t; x)$. In doing so, we assume that a thermal analysis of the problem can be made *a priori*, and in particular we neglect the effect of the cracks on the thermal conductivity of

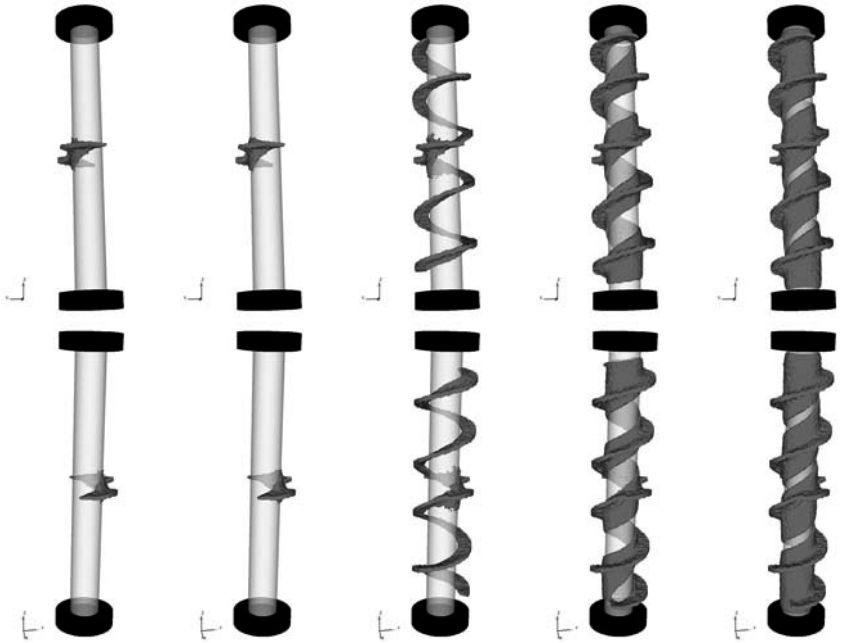


Fig. 3. Traction on a cylinder reinforced in its center. Isovolume $v \leq 5.0\text{E}-2$, representing the crack set at times t =1.5, 1.58, 1.6, 3.08 and 5.5.

the sample. Denoting by α the thermal expansion coefficient of the material, the bulk term of the total energy becomes

$$E^b(u(t^{(p)})) := \frac{1}{2}\int_\Omega A\mathbf{e}(u) : \mathbf{e}(u) - 2\frac{\alpha E}{1-\nu}\theta(t^{(p)};x)\text{tr}\mathbf{e}(u)\,dx, \qquad (11)$$

and the regularized functional E_ε becomes:

$$E_\varepsilon(u,v) := \frac{1}{2}\int_\Omega \left(v^2+\eta_\varepsilon\right)\left(A\mathbf{e}(u) : \mathbf{e}(u) - 2\frac{\alpha E}{1-\nu}\theta(t^{(p)};x)\text{tr}(\mathbf{e}(u))\right)dx + \gamma\int_\Omega \frac{(1-v)^2}{4\varepsilon} + \varepsilon|\nabla v|^2\,dx. \qquad (12)$$

The numerical implementation is similar to that described above. One major difference, is that unless temperature field depends linearly on the time (which is not the case in the experiments presented farther), the backtracking algorithm cannot be used.

Using this model, we conducted numerical simulations of a glass quenching experiments described in [26, 1, 27, 25]. We consider a thin microscope slide of width 25mm and height 75mm, with elastic moduli E=72.3GPa and $\nu = .23$ and thermal extension parameter α=7.7E-6 K^{-1}, pre-notched in its lower

end. The slide is heated up at temperature θ_H and quenched in a liquid at temperature θ_C at speed $v = 5\text{mm.s}^{-1}$. The temperature difference between the slide and the water bath is $\Delta\theta := \theta_H - \theta_C = 125\text{K}$. Neglecting the effect of the crack on the thermal properties of the sample, the temperature field in the domain is given by

$$\theta(x,y) = \begin{cases} \theta_C \text{ if } y \le vt \\ \theta_C + \Delta\theta e^{\left(-\frac{v}{\kappa}(x-vt)\right)} \text{ otherwise,} \end{cases}$$

where κ is the heat diffusion coefficient. In [25], several experiments are presented for various values of $\Delta\theta$ and v. The qualitative results are as follow, for a given $\Delta\theta$. Below a critical speed v_0, a single crack propagate along the symmetry axis. When $v_0 \le v_1$ for some v_1, the crack starts developing oscillations, then becomes unstable. For $v \ge v_1$, the behavior is more complicated and qualified of "erratic". The crack splits and each branch can in turn propagate a long a straight line, oscillate or branch. For technical reasons, we piloted our numerical experiments in terms of the fracture toughness and thermal conductivity of the material. Up to a rescaling, this is equivalent to varying the temperature difference and quenching speed. Figure 4 represents the final crack pattern for various values of γ and κ. The domain was discretized in 34,736 triangular linear elements and 17,637 nodes. The parameters ε and η_ε are respectively equal to 5E-4 and 1E-7.

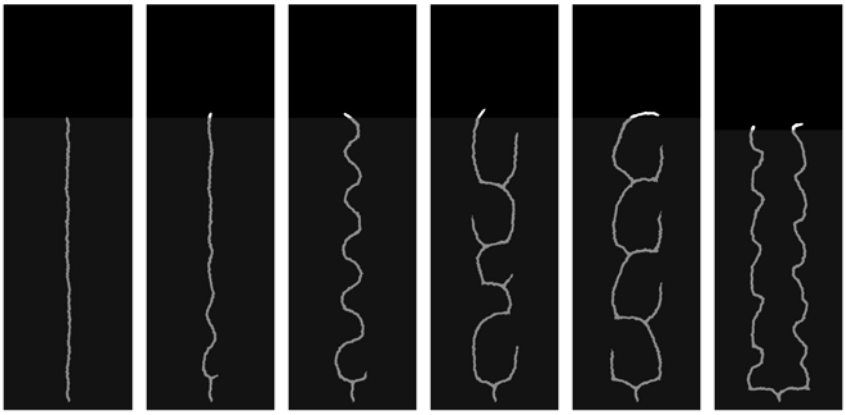


Fig. 4. Quenching of a microscope slide, the v–field for $\kappa = 17.73$ and $\gamma = 12.8, 10.0, 8.8, 7.1$ (left to right in N.m^{-1}), and $\kappa = 10$, $\gamma = 6.0$ (extreme right).

Qualitatively, the results correspond to the behavior depicted in the literature. However, in order to achieve meaningful results, we had to set κ to an unphysical high value (the real value of κ is at least two order of magnitudes lower). Using the actual value, we were only able to obtain "erratic"

propagation, unless the fracture toughness was in turn set to unrealistic high values.

Several hypothesis that could already explain this, but the problem is still open. Since the backtracking algorithm does not apply to this problem, we have no guaranty that the numerical results correspond to global minimizers. Conversely, it may be possible that the actual physical solution may be a *local* and not *global* minimizer of E_ε.

5 Conclusion and extensions

The numerical experiments presented here illustrate some the strength of the Francfort-Marigo approach, but also some pending issues with the model as well as its numerical implementations.

Beside the model's ability to predict crack path, as well as position along this path, one of the strength of Francfort and Marigo's model resides in its being applicable in two and three spaces dimensions without any modification. Indeed, none of the difficulties involved in the three–dimensional numerical implementation are related to the model itself. Instead, the major issues are the usual ones in three dimensional finite elements, *i.e.* mesh generation, visualization, file formats, or size of the discrete models, for instance.

Also, by representing the cracks in terms of the function v in (7), one avoids all difficulties related to the parameterisation of potentially complicated curves or surfaces. Note how the cracks in Figure 3 would be complicated to represent in terms of an explicit function, and how in Figure 4, simple or branched cracks are represented without special treatment. Also, again because of this representation, branching or splitting of cracks requires no special treatment to be represented.

Of course, these benefits are obtained at the cost of theoretical and numerical difficulties. In the numerical experiments and the description of the model, we had to limit the type of body forces we consider. Indeed, the formulation in its current state is inherently unable to handle surface or body forces applied to a brittle material (see the discussion in [19]). Another issue resides in the use of a Griffith-based criterion to render crack initiation and propagation simultaneously. Using cohesive energies *à la* Barenblatt may be more sound, however, the mathematical difficulties involved in doing so are tremendous and far from being solved at this point.

At the numerical level, global minimization is also very involving, as the energy to be minimized is non-convex. Future numerical work will include investigation of better minimization algorithms, as well as the study of the stability of critical points, using the Hessian of E_ε. See for that matter the analytical study of the stability of the local minimizer for a traction problem on a long beam in [8]. Lastly, numerical and mathematical evidences also suggest that in some cases, global minimization may lead to unphysical evolutions.

Alternate models involving local minimization may address that issue in the future.

Acknowledgements

Support for this work was provided by the Louisiana Board of Regents grant LEQSF (2003-06)-RD-A-05 and the National Science Foundation grant DMS-0605320. Parts of the numerical experiments were performed using the National Science Foundation TeraGrid resources provided by NCSA and PSC under the Development Allocation TG-DMS060007T and the Medium Resource Allocation TG-DMS060011N.

References

1. M. Adda-Bedia and Y. Pomeau. Crack instabilities of a heated glass strip. *Phys. Rev. E*, 52(4):4105–4113, 1995.
2. G. Alberti. Variational models for phase transitions, an approach via Γ–convergence. In G. Buttazzo *et al.*, editor, *Calculus of Variations and Partial Differential Equations*, pages 95–114. Springer–Verlag, 2000.
3. L. Ambrosio and V.M. Tortorelli. Approximation of functionals depending on jumps by elliptic functionals via Γ-convergence. *Comm. Pure Appl. Math.*, 43(8):999–1036, 1990.
4. L. Ambrosio and V.M. Tortorelli. On the approximation of free discontinuity problems. *Boll. Un. Mat. Ital. B (7)*, 6(1):105–123, 1992.
5. G. Bellettini and A. Coscia. Discrete approximation of a free discontinuity problem. *Numer. Funct. Anal. Optim.*, 15(3-4):201–224, 1994.
6. B. Bourdin. *Une méthode variationnelle en mécanique de la rupture, théorie et applications numériques*. PhD thesis, Université Paris Nord, Institut Galilée, France, 1998. available from `http://www.math.lsu.edu/~bourdin`.
7. B. Bourdin. Image segmentation with a finite element method. *M2AN Math. Model. Numer. Anal.*, 33(2):229–244, 1999.
8. B. Bourdin. Numerical implementation of a variational formulation of quasi-static brittle fracture. To appear, 08 2006.
9. B. Bourdin and A. Chambolle. Implementation of an adaptive finite-element approximation of the Mumford-Shah functional. *Numer. Math.*, 85(4):609–646, 2000.
10. B. Bourdin, G.A. Francfort, and J.-J. Marigo. Numerical experiments in revisited brittle fracture. *J. Mech. Phys. Solids*, 48(4):797–826, 2000.
11. A. Braides. *Γ-convergence for beginners*, volume 22 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2002.
12. A. Chambolle. An approximation result for special functions with bounded variations. *J. Math Pures Appl.*, 83:929–954, 2004.
13. A. Chambolle. Addendum to "an approximation result for special functions with bounded deformation" [j. math. pures appl. (9) 83 (7) (2004) 929–954]: the n-dimensional case. *J. Math Pures Appl.*, 84:137–145, 2005.

14. A. Chambolle, A. Giacomini, and M. Ponsiglione. Crack initiation in elastic bodies. To appear, 2005.
15. G. Dal Maso. *An introduction to Γ-convergence*. Birkhäuser, Boston, 1993.
16. G. Dal Maso, G.A. Francfort, and M. Toader. Quasi-static evolution in brittle fracture: the case of bounded solutions. In *Calculus of variations: topics from the mathematical heritage of E. De Giorgi*, volume 14 of *Quad. Mat.*, pages 245–266. Dept. Math., Seconda Univ. Napoli, Caserta, 2004.
17. G.A. Francfort and C. Larsen. Existence and convergence for quasi-static evolution in brittle fracture. *Comm. Pure Appl. Math.*, 56(10):1465–1500, 2003.
18. G.A. Francfort and J.-J. Marigo. Revisiting brittle fracture as an energy minimization problem. *J. Mech. Phys. Solids.*, 46(8):1319–1342, 1998.
19. G.A. Francfort and J.-J. Marigo. Griffith theory of brittle fracture revisited: merits and drawbacks. *Latin American J. Solids Structures*, 2005.
20. A. Giacomini. Ambrosio-Tortorelli approximation of quasi-static evolution of brittle fractures. *Calc. Var. Partial Differential Equations*, 22(2):129–172, 2005.
21. Alessandro Giacomini and Marcello Ponsiglione. A discontinuous finite element approximation of quasi-static growth of brittle fractures. *Numer. Funct. Anal. Optim.*, 24(7-8):813–850, 2003.
22. L. Modica and S. Mortola. Il limite nella Γ–convergenza di una famiglia di funzionali ellittici. *Boll. Un. Mat. Ital. A (5)*, 14(3):526–529, 1977.
23. L. Modica and S. Mortola. Un esempio di Γ^-–convergenza. *Boll. Un. Mat. Ital. B (5)*, 14(1):285–299, 1977.
24. M. Negri. A finite element approximation of the Griffith model in fracture mechanics. *Numer. Math.*, 95:653–687, 2003.
25. B. Yang and K. Ravi-Chandar. Crack path instabilities in a quenched glass plate. *J. Mech. Phys. Solids*, 49:91–130, 2000.
26. A. Yuse and A.M. Sano. Transition between crack patterns in quenched glass plates. *Nature*, 362:329–330, 1993.
27. A. Yuse and M. Sano. Instability of quasi-static crack patterns in quenched glass plates. *Physica D*, 108:365–378, 1997.

Measurement and Identification Techniques for Evolving Discontinuities

François Hild, Julien Réthoré, and Stéphane Roux

LMT-Cachan, ENS Cachan/CNRS-UMR 8535/Université Paris 6,
61 avenue du Président Wilson, F-94235 Cachan Cedex, France
hild@lmt.ens-cachan.fr

Summary. Digital Image Correlation allows one to estimate displacement fields based on a series of digital images of the surface of a specimen subjected to a specific loading history. Recent advances have been achieved through a novel formulation that enables one to decompose the searched displacement field onto a suited library of such fields. The latter are either finite element shape functions, which open the way to a further identification step, or mechanically significant fields such as analytic displacement fields for cracks. The interest of this experimental tool is that it provides full kinematic fields. For heterogeneous tests or discontinuities, it thus gives access to a wealth of data that are exploited to estimate mechanical properties. A first route is given by post-processing measured displacements. Two options are followed. First, by using a standard least squares technique, stress intensity factors and crack tip locations are determined by using a known displacement basis. Second, an integral interaction formulation gives also access to stress intensity factors. By choosing suitable test functions, the minimization of a scalar product with respect to measurement noise yields the optimal basis to extract stress intensity factors. An alternative route consists in using the displacement basis directly at the measurement stage. The different procedures are applied to the analysis of a high cycle fatigue experiment on a cracked sample.

Key words: Crack tip location, digital image correlation (DIC), integrated approach, interaction integral, Photomechanics, Q4 finite elements, stress intensity factor evaluation, subsingular and supersingular displacement fields.

1 Introduction

The simulation of structural failure by numerical tools, such as the finite element method [1], has passed into common usage in many different areas of industry. Most of commercial codes deliver displacements, stresses and strains, yet rarely were the codes used to predict structural failure. The reason for

Alain Combescure et al. (eds.), IUTAM Symposium on Discretization Methods for Evolving Discontinuities, 395–412.

this has been that failure is a non-linear process and non-linear codes are difficult to use and non-friendly for design engineers. This is now changing. For example, the automotive industry has been using dynamic elasto-plastic codes for evaluating crashworthiness for over a decade. The same description will simulate metal forming, and has been accepted by the airframe and engine companies, and by the certification authorities for some component tests. Virtual testing needs to establish its reliability by industrial and academic research establishments. Apart from developing the numerical tools themselves, there is also a need for validating not only the constitutive models but also the numerical procedures by comparisons with experimental data.

The present paper is devoted to the development of measurement procedures and identification techniques suitable for numerical simulations using, for instance, finite element codes. One key quantity to bridge the gap between experiments and simulations is a displacement field. Digital Image Correlation (DIC) allows one to estimate full displacement fields based on a series of digital images of the surface of a specimen subjected to a specific loading history [2, 3]. Recent advances have been achieved through a novel formulation that enables one to decompose the searched displacement field onto a suited library of such fields. The latter are either finite element shape functions [4], which open the way to a further numerical identification procedure, or mechanically significant fields such as analytic displacement fields for cracks [5], which directly provide mechanical parameters such as elastic properties [6], stress intensity factors or toughness. It is believed that DIC techniques may offer an ideal interface between experiments and numerical modeling, thus allowing for major progress in the reliability of virtual testing, and hence its future wide use and dissemination for demanding applications.

This paper is organized as follows. Displacement measurements through DIC are presented in a rather general framework, including recent advances such as the multiscale approach that endow the technique with more robustness and less uncertainty [4]. Typical performances in terms of displacement uncertainty are reached commonly in the range 10^{-2} to 10^{-3} pixel size. The interest of this experimental tool is that it provides full kinematic fields. For heterogeneous tests or discontinuities, it thus gives access to a wealth of data that is to be exploited to estimate mechanical properties. A first route is given by post-processing measured displacements. Two options are followed to analyze discontinuities. First, with a standard least squares technique, stress intensity factors and crack tip locations are determined by using a known displacement basis. Second, an integral interaction formulation gives also access to stress intensity factors. By choosing suitable test functions, measurement noise sensitivity is minimized yielding the optimal basis to extract stress intensity factors. An alternative route consists in using the displacement basis directly at the measurement stage. In such a case, referred to as *integrated* Digital Image Correlation (or I-DIC [5]), there is no decoupling between the measurement and identification stages. A fatigue crack in steel (CCT geometry) is studied and the three identification techniques are compared.

2 Principles of Image Correlation

The principle of image correlation is the following. Two digital (gray level) images corresponding respectively to a reference, $f(\mathbf{x})$, and deformed, $g(\mathbf{x})$, state are simply related by the "passive" advection of the local texture in the displacement field $\mathbf{u}$. Therefore it is assumed that

$$g(\mathbf{x}) = f(\mathbf{x} + \mathbf{u}(\mathbf{x})) \tag{1}$$

From the knowledge of f and g, the problem consists in estimating $\mathbf{u}$ as accurately as possible, usually by regularization of the conservation equation of the optical flow [7, 8, 9, 10, 11].

2.1 Classical Approach

Whenever $\mathbf{u}$ is a simple rigid translation, a powerful technique consists in looking for the maximum cross-correlation between f and g. The latter is reached accurately for the displacement $\mathbf{u}$, and suited interpolation techniques allow one to reach sub-pixel uncertainty (down to 10^{-2} pixel or below [12]). From this observation a natural scheme is to decompose the original image into small zones of interest, where the approximation of a local rigid translation or uniform strain may hold. The collocation of all the local translations thus provides an estimate of the global displacement field. Each elementary determination is performed either in the physical space [2, 3] or in Fourier space [13, 14, 15]. One difficulty however is the accurate pairing of zones of interest which implies a good evaluation of the displacement prior to further refinement. This has led to the development of multiscale algorithms that proceed by iteration from a coarse description to a finer one [16].

2.2 General Framework

The conservation of the optical flow (1) is rewritten by assuming that the reference image is differentiable and consists in minimizing the local residual Φ

$$\Phi^2(\mathbf{x}) = [\mathbf{u}(\mathbf{x}).\nabla f(\mathbf{x}) + f(\mathbf{x}) - g(\mathbf{x})]^2 \tag{2}$$

The measurement of the displacement is an ill-posed problem. The displacement is only measurable along the direction of the intensity gradient. Consequently, additional hypotheses have to be proposed to solve the problem. To estimate $\mathbf{u}$, the quadratic difference [right hand side of Eq. (2)] is integrated over the studied domain Ω and subsequently minimized

$$\eta^2 = \iint_{\Omega} [\mathbf{u}(\mathbf{x}).\nabla f(\mathbf{x}) + f(\mathbf{x}) - g(\mathbf{x})]^2 \, \mathrm{d}\mathbf{x} \tag{3}$$

The displacement field is decomposed over a set of functions $\mathbf{\Psi}_n(\mathbf{x})$. Each component of the displacement field is treated in a similar manner, and thus only *scalar* shape functions $\psi_n(\mathbf{x})$ are introduced

$$\mathbf{u}(\mathbf{x}) = \sum_{\alpha,n} a_{\alpha n} \psi_n(\mathbf{x}) \mathbf{e}_\alpha \tag{4}$$

where $\mathbf{e}_\alpha$ are elementary unit vectors along each space dimension α and $a_{\alpha n}$ the corresponding amplitudes. The objective function is thus expressed as

$$\eta^2 = \iint_\Omega \left[\sum_{\alpha,n} a_{\alpha n} \psi_n(\mathbf{x}) \nabla f(\mathbf{x}).\mathbf{e}_\alpha + f(\mathbf{x}) - g(\mathbf{x}) \right]^2 \mathrm{d}\mathbf{x} \tag{5}$$

and hence its minimization leads to a *linear* system

$$\begin{aligned} \sum_{\beta,m} a_{\beta m} \iint_\Omega [\psi_m(\mathbf{x}) \psi_n(\mathbf{x}) \partial_\alpha f(\mathbf{x}) \partial_\beta f(\mathbf{x})] \mathrm{d}\mathbf{x} = \\ = \iint [g(\mathbf{x}) - f(\mathbf{x})] \, \psi_n(\mathbf{x}) \partial_\alpha f(\mathbf{x}) \mathrm{d}\mathbf{x} \end{aligned} \tag{6}$$

that is written in a compact form as

$$\mathbf{M}\mathbf{a} = \mathbf{b} \tag{7}$$

where $\partial_\alpha f = \nabla f.\mathbf{e}_\alpha$ denotes the directional derivative, the matrix $\mathbf{M}$ and the vector $\mathbf{b}$ is directly read from Eq. (6)

$$M_{\alpha n \beta m} = \iint_\Omega [\psi_m(\mathbf{x}) \psi_n(\mathbf{x}) \partial_\alpha f(\mathbf{x}) \partial_\beta f(\mathbf{x})] \mathrm{d}\mathbf{x} \tag{8}$$

and

$$b_{\alpha n} = \iint_\Omega [g(\mathbf{x}) - f(\mathbf{x})] \, \psi_n(\mathbf{x}) \partial_\alpha f(\mathbf{x}) \mathrm{d}\mathbf{x} \tag{9}$$

The present development is similar to a Rayleigh-Ritz procedure frequently used in elastic analyses [1]. The only difference corresponds to the fact that the variational formulation is associated with the (linearized) conservation of the optical flow and not the principal of virtual work.

2.3 First Case: Q4-DIC

A large variety of functions $\mathbf{\Psi}$ may be considered. Among them, finite element shape functions are particularly attractive because of the interface they provide between the measurement of the displacement field and a numerical modeling of it based on a constitutive equation. Whatever the strategy chosen for the identification of the constitutive parameters, choosing an identical kinematic description suppresses spurious numerical noise at the comparison step. Moreover, since the image is naturally partitioned into pixels, it is appropriate to choose a square or rectangular shape for each element. This leads us to the choice of Q4-finite elements as the simplest basis. Each element is mapped onto the square $[-1, 1]^2$, where the four basic (*i.e.,* P1) functions are

$(1 \pm x)(1 \pm y)$ in a local (x, y) frame. The displacement decomposition (4) is therefore particularized to account for the previous shape functions of a finite element discretization. Each component of the displacement field is treated in a similar manner, and thus only *scalar* shape functions $N_n(\mathbf{x})$ are introduced to interpolate the displacement $\mathbf{u}^e(\mathbf{x})$ in an element Ω_e

$$\mathbf{u}^e(\mathbf{x}) = \sum_{n=1}^{n_e} \sum_{\alpha} a^e_{\alpha n} N_n(\mathbf{x}) \mathbf{e}_\alpha \tag{10}$$

where n_e is the number of nodes (here $n_e = 4$), and $a^e_{\alpha n}$ the unknown nodal displacements. The objective function is recast as

$$\eta^2 = \sum_e \iint_{\Omega_e} \left[\sum_{\alpha,n} a^e_{\alpha n} N_n(\mathbf{x}) \nabla f(\mathbf{x}).\mathbf{e}_\alpha + f(\mathbf{x}) - g(\mathbf{x}) \right]^2 \mathrm{d}\mathbf{x} \tag{11}$$

and hence its minimization leads to a *linear* system (6) in which the matrix $\mathbf{M}$ is obtained from the assembly of the elementary matrices $\mathbf{M}^e$ whose components read

$$M^e_{\alpha n \beta m} = \iint_{\Omega_e} [N_m(\mathbf{x}) N_n(\mathbf{x}) \partial_\alpha f(\mathbf{x}) \partial_\beta f(\mathbf{x})] \mathrm{d}\mathbf{x} \tag{12}$$

and the vector $\mathbf{b}$ corresponds to the assembly of the elementary vectors $\mathbf{b}^e$ such that

$$b^e_{\alpha n} = \iint_{\Omega_e} [g(\mathbf{x}) - f(\mathbf{x})] \, N_n(\mathbf{x}) \partial_\alpha f(\mathbf{x}) \mathrm{d}\mathbf{x} \tag{13}$$

Thus it is straightforward to compute for each element e the elementary contributions to $\mathbf{M}$ and $\mathbf{b}$. The latter is assembled to form the global "rigidity" matrix $\mathbf{M}$ and "force" vector $\mathbf{b}$, as in standard finite element problems [1]. The only difference is that the "rigidity" matrix and the "force" vector contain picture gradients in addition to the shape functions, and the "force" vector includes also picture differences. The matrix $\mathbf{M}$ is symmetric, positive and sparse. These properties are exploited to solve the linear system efficiently. Moreover, the only parameter the user has to choose is the size ℓ of the considered elements. A regular mesh is used and is made of squares containing $\ell \times \ell$ pixels. The size ℓ has a direct impact on the measurement uncertainty as will be shown later on. Last, the domain integrals involved in the expression of $\mathbf{M}^e$ and $\mathbf{b}^e$ require imperatively a pixel summation. The classical quadrature formulas (*e.g.,* Gauss point) is *not* used because of the very irregular nature of the image texture. This latter property is crucial to obtain an accurate displacement evaluation. Furthermore, sub-pixel interpolation is used to decrease the resolution of the algorithm. The very use of a Taylor expansion requires that the displacement be small when compared with the correlation length of the texture. For a fine texture and a large initial displacement, this requirement appears as inappropriate to converge to a meaningful solution.

Thus one may devise a generalization to arbitrarily expand the correlation length of the texture. This is achieved through a coarse-graining step. The interested reader may refer to Ref. [4] for additional details concerning the above mentioned features.

2.4 Second Case: I-DIC

The route outlined in Sect. 2.2 is still followed, but the relevant displacement fields, $\boldsymbol{\Psi}_i$, are introduced *a priori* at the measurement stage. It is referred to as an *integrated* approach of measurement and identification, or I-DIC [5]. The direct computation of the amplitudes of the singular crack fields will thus immediately provide an estimate for the SIFs. The generality of the correlation approach is fully exploited, and no specific restrictions appear from having to deal with eight functions. First, rigid body motions have to be accounted for, and constitute the first three fields, two in-plane translations, $\boldsymbol{\Psi}_1$ and $\boldsymbol{\Psi}_2$, and one rotation about an axis normal to the observation plane, $\boldsymbol{\Psi}_3$. One single uniform strain field, $\boldsymbol{\Psi}_4$, is allowed for in order to leave the crack face stress-free, the so-called "T-stress" (uniaxial tension along the crack direction). The presence of the crack generates two independent singular strain fields corresponding to modes I, $\boldsymbol{\Psi}_5$, and II, $\boldsymbol{\Psi}_6$, with two scalar amplitudes that will give access to the stress intensity factors. Last, in order to enrich the basis, one may also consider the next order sub-singular mode I and II fields, respectively denoted by $\boldsymbol{\Psi}_7$ and $\boldsymbol{\Psi}_8$. It is very convenient to express these vectors as complex valued fields in the local crack frame (crack tip at the origin, and crack extension along the negative real axis). A current point M is represented by a complex number $\mathbf{z} = x + iy$ or polar coordinates $\mathbf{z} = re^{i\theta}$, and similarly the displacement is represented by $\mathbf{u} = u_x + iu_y$. The expression of the eight basis functions reads

$$\begin{aligned}
\boldsymbol{\Psi}_1(\mathbf{z}) &= 1 \\
\boldsymbol{\Psi}_2(\mathbf{z}) &= i \\
\boldsymbol{\Psi}_3(\mathbf{z}) &= iz \\
\boldsymbol{\Psi}_4(\mathbf{z}) &= (\kappa - 1)z + 2\overline{z} \\
\boldsymbol{\Psi}_5(\mathbf{z}) &= \sqrt{r}[2\kappa e^{i\theta/2} - e^{3i\theta/2} - e^{-i\theta/2}] \\
\boldsymbol{\Psi}_6(\mathbf{z}) &= i\sqrt{r}[2\kappa e^{i\theta/2} + e^{3i\theta/2} - 3e^{-i\theta/2}] \\
\boldsymbol{\Psi}_7(\mathbf{z}) &= \sqrt{r^3}[2\kappa e^{3i\theta/2} - 3e^{i\theta/2} + e^{-3i\theta/2}] \\
\boldsymbol{\Psi}_8(\mathbf{z}) &= i\sqrt{r^3}[2\kappa e^{3i\theta/2} + 3e^{i\theta/2} - 5e^{-3i\theta/2}]
\end{aligned} \tag{14}$$

where

$$\kappa = \frac{(3-\nu)}{(1+\nu)} \tag{15}$$

in plane stress condition, as expected along the free observation surface, with ν being the Poisson's ratio. Let us note however that the amplitudes of these functions are real numbers. The motivation for choosing such a basis will be described further down.

3 Stress Intensity Factor Identification

3.1 First Method: Least Squares Minimization

Extracting some mechanically meaningful information using the detailed map of displacement is performed by identifying the amplitudes of relevant reference displacement fields. In the present case, one lists easily meaningful contributions [see Eq. (14)]. The strategy is thus to decompose the estimated displacement field (from image correlation) onto the basis of $\mathbf{\Psi}_k$ test functions. For this goal, the following objective function is minimized

$$\mathcal{T}(\mathbf{a}, \mathbf{z}_0, \theta_0) = \sum_{n=1}^{N_{\text{mes}}} |\mathbf{u}_n[(\mathbf{z}_n - \mathbf{z}_0)e^{-i\theta_0}] - \sum_k a_k \mathbf{\Psi}_k(\mathbf{z}_n)|^2 \tag{16}$$

where N_{mes} is the number of measured displacement points, a_k the sought amplitudes, $\mathbf{z}_n$ the location of the n-th measurement point, $\mathbf{z}_0$ the location of the crack tip in the original frame, and θ_0 the orientation of the crack front.

One advantage of this formulation is that the value reached by the objective function constitutes a global quality parameter. Since the identification procedure assumes that the crack geometry is known, a minimization of this global residual over the *a priori* guessed crack tip position also provides a natural way of optimizing the crack geometry.

3.2 Second Method: Interaction Integral

It is worth noting that other techniques are followed to extract the stress intensity factor by computing the crack opening displacement profile [17] or by using an interaction integral formalism [18, 19, 20]. The latter is now discussed and was shown to be applicable to experimental tests [21].

Let us introduce a displacement field $\mathbf{u}$ and the corresponding stress field $\boldsymbol{\sigma}$, which are solutions of a linear elastic problem with known elastic constants, and whose geometry contains a linear crack. To measure the stress intensity factor at the crack tip, it is convenient to use auxiliary displacement and stress fields, $\mathbf{u}^{aux}$ and $\boldsymbol{\sigma}^{aux}$, (related through the same constitutive equation and domain geometry) and the so-called virtual crack extension (or VCE in the sequel) field $\mathbf{q}$. The interaction integral I^{int} is defined as

$$I^{int} = -\int_\Omega \left[\sigma^{aux}_{ml} u_{m,l} \delta_{kj} - (\sigma^{aux}_{ij} u_{i,k} + \sigma_{ij} u^{aux}_{i,k})\right] q_{k,j} \, \mathrm{d}\Omega \tag{17}$$

where δ_{kj} is the Kronecker operator. In this expression, the VCE field $\mathbf{q}$ simply appears as a (vector) weight function. It is assigned a value $\mathbf{x}$ at the tip and $\mathbf{0}$ on the boundary of the integration domain $\partial\Omega$. It is also parallel to the crack faces

$$\begin{aligned} \mathbf{q} \cdot \mathbf{n}_\Gamma &= 0 \quad \text{on } \Gamma \\ \mathbf{q} &= \mathbf{x} \quad \text{at the tip} \\ \mathbf{q} &= \mathbf{0} \quad \text{on } \partial\Omega \end{aligned} \tag{18}$$

where Γ is the geometrical support of the crack and $\mathbf{n}_\Gamma$ is the normal to Γ. It is shown theoretically that I^{int} is domain-independent and that the interaction integral value only depends on the stress intensity factors of the actual and auxiliary fields [22], or more precisely in plane stress

$$I^{int} = \frac{2}{E}\left(K_I K_I^{aux} + K_{II} K_{II}^{aux}\right) \tag{19}$$

where K_i, respectively K_i^{aux}, are the stress intensity factors of the actual, respectively auxiliary, field, and E the Young's modulus. Consequently, one computes K_I (respectively K_{II}) choosing the auxiliary field to be the Westergaard solution for a cracked body in mode I (resp. mode II) with a unit stress intensity factor. This technique has been widely used in numerical simulations with a VCE field $\mathbf{q}$, which is everywhere parallel to the crack tip, and whose modulus varies linearly with the distance to the crack tip from 1 at the tip, to 0 at the edge of a circular domain (this geometry is referred to as "conical" in the sequel).

For identification purposes, the measured displacement field is chosen for $\mathbf{u}$, while $\boldsymbol{\sigma}$ is estimated from the displacement field and the constitutive equation, so that the interaction integral is a linear operator acting on the measured displacement field. A difficulty is that the interaction integral involves the gradient of the measured displacement field and singular auxiliary fields. Therefore, a high noise sensitivity of the stress intensity factors estimation by this technique is expected (and observed). However, only the satisfaction of the boundary conditions [Eq. (18)] is required for the result (19) to hold, and hence the VCE field is adjusted to reduce the noise sensitivity of the stress intensity factors identification.

A discretization scheme similar to that of the displacement is written for $\mathbf{q}$

$$\mathbf{q}^e(\mathbf{x}) = \sum_{n=1}^{n_e} \sum_{\alpha} q_{\alpha n}^e N_n(\mathbf{x}) \mathbf{e}_\alpha \tag{20}$$

Inserting the discrete displacement and virtual crack extension into the interaction integral I^{int}, one obtains a generic stress intensity factor K (for mode I or mode II) in a matrix / vector form, where $\mathbf{Q}$ (respectively $\mathbf{A}$) collect the VCE nodal degrees of freedom (resp. the displacement)

$$K = \mathbf{Q}^{\mathrm{T}} \mathbf{L} \mathbf{A} \tag{21}$$

In this equation, K and $\mathbf{L}$ are specialized for mode I or mode II by selecting the appropriate auxiliary field.

Let us now assume that the actual approximated displacement field is corrupted by a random (vector) noise $\boldsymbol{\zeta}$ of zero mean

$$\langle \boldsymbol{\zeta}(\mathbf{x}) \rangle = \mathbf{0} \tag{22}$$

This noise is also characterized by its variance γ^2 and more completely by its (normalized) correlation kernel $\mathbf{C}$, so that the auto-correlation function is written as

$$\langle \boldsymbol{\zeta}(\mathbf{x})\boldsymbol{\zeta}(\mathbf{x}+\mathbf{y})\rangle_{\mathbf{x}} = \gamma^2 \mathbf{C}(\mathbf{y}) \tag{23}$$

For an uncorrelated isotropic white noise, $C_{ij}(\mathbf{x}) = \delta_{ij}\delta(\mathbf{x})$ and the variance of the noise along any direction is γ^2.

The measurement uncertainties affect the estimation of the stress intensity factors. Using the linearity of Eq. (21), the perturbation δK induced by noise $\boldsymbol{\zeta}$ is of zero mean

$$\langle \delta K \rangle = \mathbf{Q}^{\mathrm{T}}\mathbf{L}\ \langle \boldsymbol{\zeta} \rangle = 0 \tag{24}$$

and its variance reads

$$\langle \delta K^2 \rangle = \mathbf{Q}^{\mathrm{T}}\ \mathbf{L}\mathbf{C}\mathbf{L}^{\mathrm{T}}\ \mathbf{Q}\ \gamma^2 \tag{25}$$

From this equation, the variance $\langle \delta K^2 \rangle$ depends upon the nodal values of the VCE field, both displacement and weight functional bases, the auxiliary field and the noise kernel. However, it is independent of the measured displacement field. An interesting feature of Eq. (25) is that the K variance is a quadratic form of $\mathbf{Q}$. Hence, it is exploited for a minimization under the constraints given by the boundary conditions (18). The optimal weight function for mode I is displayed in Fig. 1.

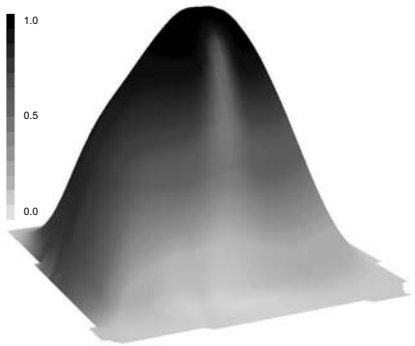


Fig. 1. Shape of the optimal VCE computed for mode I.

Figure 2 illustrates the noise sensitivity reduction obtained by the optimization of the VCE field. It is shown that the variance $\langle \delta K^2 \rangle$ for a conical VCE is decreasing as a^{-2} where a is the size of the integration domain. For the optimized VCE, the decay is steeper, scaling as a^{-3}. Consequently, one predicts that the ratio of the noise reduction from the cone to the optimum VCE weight decreases with the domain size as $1/a$. For typical sizes that are considered in digital image correlation applications, the decrease in variance reaches up to two orders of magnitude. More details about the improvements brought out by the VCE optimization are given in Ref. [23].

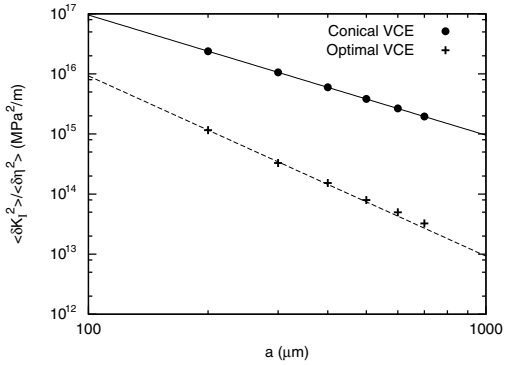


Fig. 2. Noise sensitivity for mode I with the cone and the optimal weight functions.

3.3 Third Method: I-DIC

Up to now, the identification procedures discussed herein consisted of two *independent* steps:

- first, the measurement of the displacement field, and in particular the use of an arbitrary functional basis for that purpose,
- second, the projection of the measured displacement field onto a few mechanically significant fields.

It is worth noting that the separation into two steps is unnecessary, and in particular the introduction of Q4-elements is artificial, and only used at an intermediate stage. This *a priori* partition will reduce the performance of the method, as compared to an integrated approach. This third method has already been presented in Section 2.4. Since the trial fields already contain the pure mode I and mode II cases, the respective stress intensity factors are directly the amplitude of those fields, and no further treatment is necessary.

3.4 Comparison between the different approaches

Since three different methods have been presented, it is worth discussing their expected merits and weaknesses. Let us first note that the three approaches require the knowledge of the elastic properties of the uncracked medium. If they are not known or spatially varying the method is not directly applicable. Furthermore, the present approaches are restricted to quasi-static situations since inertia terms that arise in dynamic fracture are not included. This is however not an intrinsic limitation of the method, since this generalization may be performed in the same spirit as the quasi-static case. As mentioned earlier, there is a major difference between the first two methods, and the last one, because of the additional projection of the kinematic field onto a functional basis that does not contain any specificities inherited from the

analysis of the mechanical problem itself. The first two methods will thus be compared first.

Starting for a discretized kinematic field, both routes 1 and 2 aim at projecting the field onto the pure crack singularity. However, this projection is based on a different metric. In the first case, the Euclidian scalar product of the actual and reference field is used, whereas in the second case, the scalar product is traced back as the elastic "co-energy" between actual and auxiliary fields. Therefore, one sees the advantage of the second method, in the sense that the projection will be orthogonal to all (other than the auxiliary) elastic displacement fields on the same geometry (but potentially with different boundary conditions remote from the crack tip). Therefore, the presence of a boundary, a complex (or even unknown) loading, will not affect the determination of the stress intensity factor. Moreover, a slight error in the positioning of the crack tip is investigated using a perturbation approach. In this case, on top of the standard singularity, a stronger one will appear (displacements decaying as $1/\sqrt{r}$ away from the crack tip), termed generically super-singular field. By construction, the latter is naturally the solution to an elastic problem on the same geometry, and hence is orthogonal (with the elastic co-energy metric) to the reference field. Consequently, a slight error in the crack tip location along its axis should have no impact on the measurement of K, and hence this estimate should be quite robust. Let us however underline that as elasticity is an essential ingredient included in the chosen metric, crack tip plasticity, damage, or any other non-linear behavior may affect this property.

This property is not valid for the first approach. Yet, it is artificially reintroduced. It suffices to enrich the basis of trial fields with well chosen functions to limit their influence. In particular, supersingular fields are introduced in the library of basis functions. Subsingular fields are also considered as $\boldsymbol{\Psi}_7$ and $\boldsymbol{\Psi}_8$ in Eq. (14), to account for complex loadings. The limitation is that such fields have to be explicitly constructed. There is however an advantage of this method, since the amplitude of all fields will be estimated from the projection. For instance, the amplitude of the first supersingular field will give a very precise estimate of the ideal location of the crack tip. The other property of the first approach is that it is based on the minimization of a functional, and thus minimization with respect to different parameters are considered, such as the crack geometry, or Poisson's ratio, to extract this information. Yet another interesting feature, illustrated below, is the fact that this minimization is the surface integral of an "error" density. Thus the latter field itself is extremely informative, namely, it indicates whether systematic deviations are observed close to the crack tip (*e.g.,* small scale yielding) or remote from it (*e.g.,* presence of a boundary), and thus the field library are enriched or the domain limited to account for this. It may also reveal specific regions in the sample where the presence of defects are detected (*e.g.,* porosity, micro-crack) and that are to be excluded from the analyzed kinematic field.

Most of the characteristics of the first method also hold for the third one, with the additional benefit of dealing with the kinematic field precisely at the

pixel level, and not after a first filtering due to the kinematic analysis functions (such as Q4 finite elements used herein). After this general discussion, the application of the three approaches to a real experimental test is now presented.

4 Analysis of a Fatigue Experiment

4.1 Experimental Configuration and Kinematic Measurements

The studied material is an XC48 (or C45) steel with a Young's modulus of 190 GPa, a Poisson's ratio of 0.3. The cyclic yield strength (offset: 0.2 %) is equal to 210 MPa. The sample has a CCT geometry (Fig. 3-a) and is subjected to cyclic tension with a load ratio $R = 0.4$. In the present analysis, only the stage corresponding to the maximum load level is considered after about 300,000 cycles for which the crack size $2a = 14.5$ mm. Pictures are taken by using a long distance microscope and a CCD camera (resolution: 1024×1280 pixels, dynamic range: 12 bits) so that the physical size of one pixel is 2.08 μm. At this magnification, the raw surface is observed (Fig. 3b). Contrary to other materials (*e.g.,* aluminium alloy [24]) no special surface preparation was used.

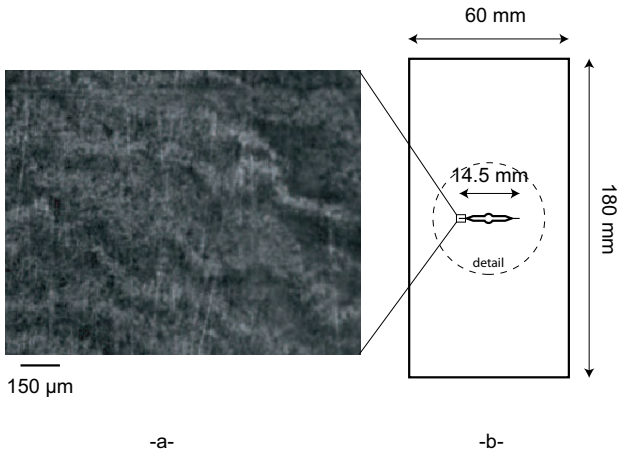


Fig. 3. -a-Reference picture of the CCT test (12-bit digitization, 1024×1280 pixel resolution). -b-CCT geometry.

The displacement field is measured by using the Q4-DIC algorithm (see Sect. 2.3). Prior to any actual analysis, it is important to estimate the performance of the approach on the texture of the image (Fig. 3-a). If one changes the picture, one may not get exactly the same performance since it is related

to the local details of the gray level distribution as shown in Sect. 2.3. This is performed by using the original image f only, and generating a translated image g by a prescribed amount $\mathbf{u}_{pre}$. Such an image is generated in Fourier space using a simple phase shift for each amplitude [25]. This procedure implies a specific interpolation procedure for inter-pixel gray levels, to which one resorts systematically [4]. The algorithm is then run on the pair of images (f, g), and the estimated displacement field $\mathbf{u}_{est}(\mathbf{x})$ is measured. One is mainly interested in sub-pixel displacements, where the main origin of errors comes from inter-pixel interpolation. Therefore the prescribed displacement is chosen along the $(1, 1)$ direction so as to maximize the interpolation sensitivity. To highlight this reference to the pixel scale, one refers to the x- (or y-) component of the displacement $u_{pre} \equiv \mathbf{u}_{pre}.\mathbf{e}_x$ varying from 0 to 1 pixel, rather than the Euclidian norm (varying from 0 to $\sqrt{2}$ pixels).

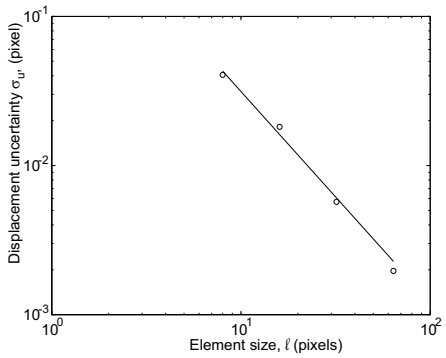


Fig. 4. Displacement uncertainty vs. element size for the picture shown in Fig. 3 with a Q4-DIC algorithm.

The quality of the estimate is characterized by the standard uncertainty $\sigma_u = \langle \|\mathbf{u}_{est} - \langle \mathbf{u}_{est} \rangle\|^2 \rangle^{1/2}$. It reaches a maximum for one half pixel displacement, $u_{pre} = 0.5$ pixel, and is approximately symmetric about this maximum [4]. Integer valued displacements (in pixels) imply no interpolation and are exactly captured through the multi-scale procedure discussed above. To quantify the effect of the element size ℓ, the standard uncertainty is averaged over u_{pre} within the range $[0, 1]$ as a function of the element size ℓ. The result is shown in Fig. 4. A power-law decrease

$$\langle \sigma_u \rangle = A^{1+\zeta} \ell^{-\zeta} \tag{26}$$

for $8 \leq \ell \leq 64$ pixels is observed as shown by a line on the graph. The amplitude is close to 1 pixel (more precisely $A = 0.9$ pixel). The exponent is measured to be $\zeta = 1.4$. These results quantify the trade-off the experimentalist has to face in the analysis of a displacement field, namely, either the measurement is accurate but estimated over a large zone, or it is spatially

resolved but at the cost of a less accurate determination. This is a significant difference with classical finite element techniques for which convergence is achieved when the element size decreases [1].

Even though the displacement uncertainty is the lowest for large element sizes, the heterogeneity of the displacement prompts one to use small sizes. To achieve a displacement uncertainty less than 2.5×10^{-2} pixel (or 52 nm), elements of size 12×12 pixels.

4.2 Analysis of the Maximum Load Level

This study aims at comparing the three discussed methods to identify stress intensity factors. It is performed for the maximum load level of the cycle. Figure 5 shows a map of the two components of the displacement field, where the discontinuity of displacements appears clearly, mainly in mode I, although the crack does not propagate along the horizontal direction, had the experimental conditions been perfect. Consequently, an identification based on "perfect" boundary conditions, which is usually assumed by using finite element simulations, is not secure in the present case. In the analyzed configuration, there are $N_{\text{mes}} = 10788$ nodes.

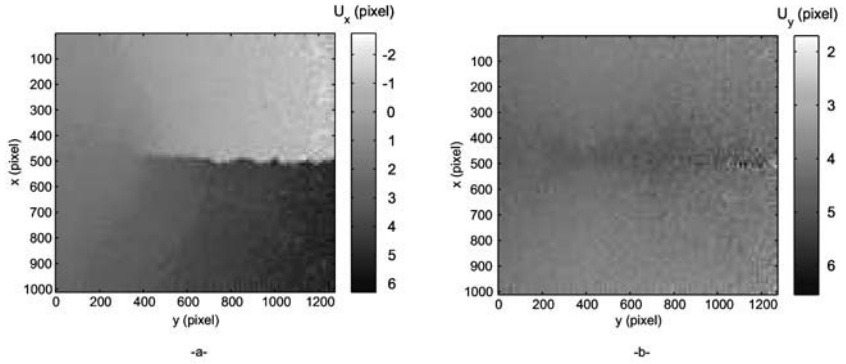


Fig. 5. Vertical (a) and horizontal (b) displacement fields expressed in pixels (1 pixel is equal to 2.08 μm) in a CCT experiment measured by using a Q4-procedure with an element size $\ell = 12$ pixels.

For any of the three identification routes, one key quantity in the error $\mathcal{T}$ or η is the location of the crack tip in the reference picture, and the angle of the crack face wrt. a reference direction in the picture. For each given location of the crack, the values of the amplitudes of the fields $\mathbf{\Psi}_k$ is computed by minimizing $\mathcal{T}$ or η.

When using the first identification route, it is shown that there is no need to account for mode II terms (*i.e.,* $\mathbf{\Psi}_6$ and $\mathbf{\Psi}_8$ are discarded) since the values

of the stress intensity factor K_{II} is less than 3 % that in mode I and the influence on the identification error is vanishingly small.

Figure 6 shows a map of the two components of the identified displacement field in very good agreement with the measured data (Fig. 5). A value of $K_I = 20$ MPa$\sqrt{\text{m}}$ with an uncertainty on the order of 1.5 MPa$\sqrt{\text{m}}$ is obtained with the least squares minimization, when the crack tip location ($\mathbf{z}_0$, θ_0) is optimized.

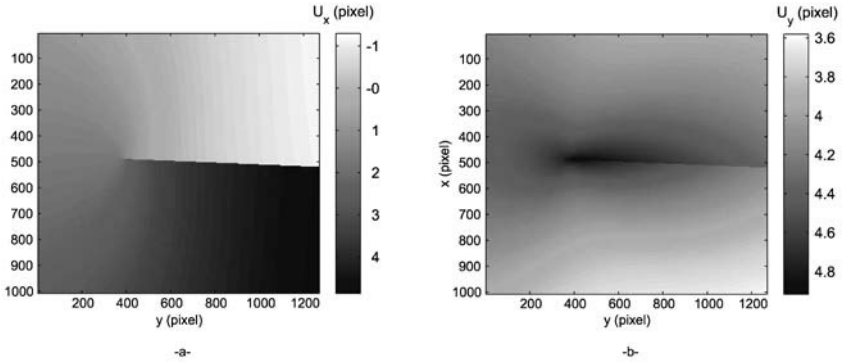


Fig. 6. Vertical (a) and horizontal (b) displacement fields expressed in pixels (1 pixel is equal to 2.08 μm) in a CCT experiment identified by using 6 displacement fields when $\ell = 12$ pixels and excluding from the analyzed a disk centered on the crack tip and of radius $R_i = 75$ pixels.

Second, the interaction integral is used with an optimized virtual crack extension field. The crack tip position is set at the same location as for the first identification technique. Figure 7 shows the results obtained with this identification route. Stress intensity factors are plotted as functions of the size of the integration domain a. The domain independence of this approach is illustrated. The quantitative estimation of the asymptotic value of K_I is 23 MPa$\sqrt{\text{m}}$ with an uncertainty on the order of 1 MPa$\sqrt{\text{m}}$. For K_{II}, the asymptotic value is extremely close to 0 MPa$\sqrt{\text{m}}$, in good agreement with the results obtained using the first method.

Last, the analysis is performed with the integrated approach. Only one element of size 512×512 pixels (or 1.13 mm^2) is used. All the pixels are considered apart those such that $\Re(\mathbf{z}) < 0$ and $\Im(\mathbf{z}) < 12$ pixels (to be consistent with the first a posteriori analysis), as well as $|\mathbf{z}| < R_i = 75$ pixels (or 156 μm). The correlation procedure to estimate the amplitude of the displacement field is quite similar to that presented in Ref. [4]. First, a gross determination of the displacement field is searched for using coarse-grained images, where super pixels are defined as averages of the true pixel values on squares of size $2^n \times 2^n$ pixels with $n = 3$ to start off with. After a first determination of the mean displacement, the strained image is corrected for using

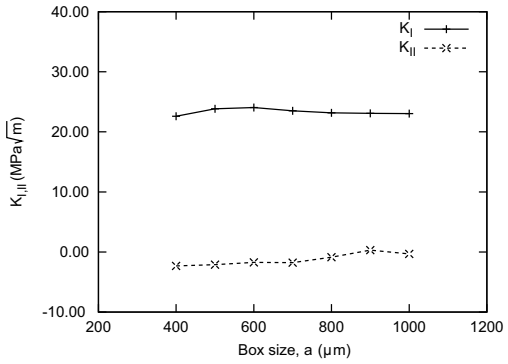


Fig. 7. Stress intensity factors from the interaction integral with an optimized VCE field, a is the size of the integration domain.

a linear interpolation. The process is repeated until convergence. Then a similar determination is carried out on a finer image where the coarse-graining is performed for $n = 2$, and similarly down to the original images. This provides a faster convergence, but more importantly, it avoids spurious trapping into secondary minima when the displacement amplitude is large. Each determination at any scale is performed using the above formulation (6) for assembling the linear system. First, the most likely crack tip position is determined. A 25 pixel (or 52 μm) difference is observed when compared with that given by the first a posteriori analysis. This may be due to the discretization error introduced with the kinematic field determination projected onto finite elements.

In the present case, the unresolved differences are measured in terms of gray levels [see Eq. (5)]. When the map of Φ is analyzed and related to the dynamic range of the considered region of interest, a maximum value of the order of 1% is obtained (Fig. 8). It follows that the displacement evaluation is deemed secure and the corresponding SIF value are trustworthy. Many unresolved differences are masked, in particular along the crack mouth. The remaining unresolved and unmasked differences are uncorrelated, thereby indicating that the basis of considered displacements is sufficient to describe the present experiment. The map of displacements is very close to those of Fig. 6. A value of $K_I = 18$ MPa$\sqrt{\text{m}}$ with an uncertainty on the order of 1 MPa$\sqrt{\text{m}}$ is obtained. Within the uncertainty levels, the three values are in a good agreement.

5 Summary

Two approaches were presented to measure a displacement field based on the comparison of two digital images. First, the sought displacement field is decomposed onto a basis of continuous functions using P1-shape functions with

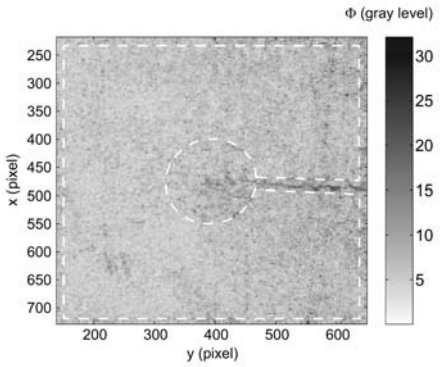


Fig. 8. Map of the residuals in gray levels at convergence for the integrated approach with $\ell = 512$ pixels. The dynamic range of the region of interest is 3052 gray levels. The dashed lines depict the masked areas.

Q4 elements as proposed in finite element methods. The latter corresponds to one of the simplest kinematic descriptions. It therefore allows for a compatibility of the kinematic hypotheses made during the measurement stage and the subsequent identification / validation stage, for instance, by using conventional finite element techniques. Based on the latter kinematic analysis, again two identification techniques for the stress intensity factors were proposed. Both of them are seen as a projection on pure elastic singular crack fields, but with different metrics. The merits of these two approaches were discussed. Second, an integrated kinematic measurement technique including an elastic identification step based on a suited basis of functions was introduced.

All those procedures were tested on an experimental test, and provided consistent estimates of mode I and mode II stress intensity factors. This paves the way to stress intensity factor measurements, and even quantitative threshold stress intensity factor or toughness estimates in situations that could not have been analyzed otherwise. Moreover, such estimates only require at most a few minutes of computation on a standard PC, and provide a global quality criterion as well as spatial error maps enabling one to adjust the analysis tool to meet any specific bias revealed therein. This study illustrates the level of accuracy that is obtained on the stress intensity factor and crack geometry using a high quality long-distance microscope, but otherwise quite common equipment. The fact that only one camera was used limits the analysis to planar surfaces. Investigating three dimensional fatigue cracks using X-ray tomography constitutes therefore a challenging direction for future investigations.

Acknowledgements

This work was supported by the CETIM Foundation grant entitled PROPAVANFIS: "Advanced methods for the experimental and numerical analysis of crack propagations under complex loadings."

References

1. Zienkievicz OC, Taylor RL (1989) The Finite Element Method. McGraw-Hill, London (UK)
2. Sutton MA, Wolters WJ, Peters WH, Ranson WF, McNeill SR (1983) Im. Vis. Comp. 1:133–139
3. Chu TC, Ranson WF, Sutton MA, Petters WH (1985) Exp. Mech. 3:232–244
4. Besnard G, Hild F, Roux S (2006) Exp. Mech. (in press)
5. Hild F, Roux S (2006) C. R. Mecanique 334:8–12. See also Roux S, Hild F (2006) Int. J. Fract. 140:141–157
6. Hild F, Roux S (2006), Strain 42:69–80
7. Horn BKP, Schunck BG (1981) Artificial Intelligence 17:185–203
8. Mitiche A, Bouthemy P (1996) Int. J. Comp. Vision 19:29–55
9. Black M (1992) Robust Incremental Optical Flow. PhD dissertation, Yale University
10. Odobez J-M, Bouthemy P (1995) J. Visual Comm. Image Repres. 6:348–365
11. Bogen D, Rahdert D (1996) IEEE Trans. Pattern Analysis and Machine Intelligence 18:629–635
12. Sutton MA, McNeill SR, Jang J, Babai M (1988) Opt. Eng. 27:870–877
13. Chen DJ, Chiang FP, Tan YS, Don HS (1993) Appl. Opt. 32:1839–1849
14. Berthaud Y, Scholz J, Thesing J (1996) Méthodes optiques et acoustiques de mesures des caractéristiques mécaniques. In: Proc. Colloque national MECAMAT "Mécanismes et mécanique des grandes déformations"
15. Chiang FP, Wang Q, Lehman F (1997) New Developments in Full-Field Strain Measurements Using Speckles. In: Non-Traditional Methods of Sensing Stress, Strain and Damage in Materials and Structures. STP 1318. ASTM, Philadelphia (USA)
16. Hild F, Raka B, Baudequin M, Roux S, Cantelaube F (2002) Appl. Optics IP 41:6815–6828
17. Forquin P, Rota L, Charles Y, Hild F (2004) Int. J. Fract. 125:171–187
18. Hellen TK (1975) Int. J. Num. Meth. Eng. 9:187–207
19. Parks DM (1977) Comp. Meth. Appl. Mech. Eng. 12:353–364.
20. Destuynder P, Djaoua M (1981) Math. Meth. Appl. Sci. 3:70–87
21. Réthoré J, Gravouil A, Morestin F, Combescure A (2005) Int. J. Fract. 132:65–79
22. Bui HD (1983) J. Mech. Phys. Solids 31:439–448
23. Réthoré J, Roux S, Hild F (2006) submitted for publication
24. Sutton MA, Zhao W, McNeill SR, Helm JD, Piascik RS, Riddel WT (1999) Local crack closure measurements: Development of a measurement system using computer vision and a far-field microscope. In: Advances in fatigue crack closure measurement and analysis: second volume. STP 1343. ASTM, Philadelphia (USA)
25. Périé JN, Calloch S, Cluzel C, Hild F (2002) Exp. Mech. 42:318–328

Conservation under Incompatibility for Fluid-Solid-Interaction Problems: the NPCL Method

E.H. van Brummelen and R. de Borst

Delft University of Technology, Faculty of Aerospace Engineering, Engineering Mechanics Chair, P.O. Box 5058, 2600 GB Delft, The Netherlands
e.h.vanbrummelen@tudelft.nl, r.deborst@tudelft.nl

Summary. Finite-element discretizations of fluid-solid-interaction problems only trivially preserve the conservation properties of the underlying problem under restrictive compatibility conditions on the approximation spaces for the fluid and the solid. The present work introduces a new general method for enforcing interface conditions that maintains the conservation properties under incompatibility. The method is based on a nonlinear variational projection of the velocity field to impose the kinematic condition, and a consistent evaluation of the load functional that accounts for the dynamic condition. Numerical results for a projection problem are presented to illustrate the properties of the method.

Key words: fluid-solid interaction, incompatibility, conservation, space-time finite-element methods.

1 Introduction

The numerical solution of fluid-solid-interaction problems has prominence in many scientific and engineering disciplines. The interaction is induced by interface conditions, which prescribe continuity of displacements and tractions across the fluid-solid interface. If the approximation spaces for the fluid and the solid in the discretization are compatible, i.e., if the fluid and the solid have identical meshes and orders of approximation at the interface, then the enforcement of these continuity requirements is trivial. However, in many instances, it is necessary to allow for incompatible approximation spaces. For instance, the meshes for the fluid and solid subsystems may have been generated by different analysts. Moreover, the disparate regularity properties of the fluid and solid solutions typically prompt distinct approximation spaces.

Alain Combescure et al. (eds.), IUTAM Symposium on Discretization Methods for Evolving Discontinuities, 413–432.

An important characteristic of fluid-solid-interaction problems pertains to their conservation properties: on account of the continuity of tractions and displacements, mass, momentum and energy are conserved at the interface and, accordingly, the interface does not appear in the conservation statements for the aggregated system. However, incompatibility impedes continuity of tractions and displacements across the interface in the discrete approximation. Consequently, incompatible finite-element discretizations of fluid-solid-interaction problems do not generally preserve the conservation properties of the underlying continuum problem.

Current coupling strategies for fluid-solid interaction are in general non-conservative. The change in a conserved quantities in the interior of the fluid and solid domains via the interface can be expressed as an inner product on the interface. Conservation requires that this inner products evaluates to the same value at both sides of the interface. Most coupling methods however fail to identify the inner products. The methodology presented in [3] identifies the inner products, but the inner products do not properly represent the change in the interior of the domains.

In this work we present a new general coupling method for fluid-solid-interaction problems that preserves the conservation properties under incompatibility. The method comprises three complementary primitives: a suitable nonlinear variational projection to impose the kinematic condition, representation of the load functional in the velocity trace space of the fluid, and a consistent evaluation of the load functional to account for the dynamic condition. We refer to the approach concisely as the NPCL (Nonlinear variational Projection with Consistent Loading) method.

2 Problem Statement

In this section, we present the governing equations for the fluid and solid subsystems and the interface conditions. Based on these equations, we then derive the conservation properties of the aggregated fluid-solid-interaction system.

2.1 Problem Setting

To furnish a setting for the model problem, let X be a Euclidean space of dimension d ($d = 2, 3$) with measure $|\cdot|$. Upon introducing orthonormal base vectors $\boldsymbol{e}_i$, $i = 1, \ldots, d$ in X, we can specify any element $\boldsymbol{x} \in X$ by its coordinates $(x_1, \ldots, x_d)$ with respect to the basis according to $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_d\boldsymbol{e}_d$. Thus, we obtain an isomorphism between X and $\mathbb{R}^d$, viz., $|\boldsymbol{x}| = \sqrt{x_1^2 + \cdots + x_d^2}$, and we can identify X and $\mathbb{R}^d$.

To accommodate the fluid and solid subproblems, we consider an open, bounded time interval $]0, T[$ and two time-dependent open, bounded subsets of X, viz., $\Omega(t)$ and $\Xi(t)$. The mappings $t \mapsto \Omega(t)$ and $t \mapsto \Xi(t)$ are continuous, so that $\Omega(t)$ and $\Xi(t)$ are well-defined for all $t \in [0, T]$. The fluid

and solid subproblems are set on $\Omega(t)$ and $\Xi(t)$, respectively. We assume that $\Omega(t)$ and $\Xi(t)$ are contiguous. This implies that the intersection of the domains is empty, i.e., the domains are non-overlapping, and the intersection of their closures is non-empty. The latter implies that there exists an *interface* $\Gamma(t) = \mathrm{cl}(\Omega(t)) \cap \mathrm{cl}(\Xi(t)) \neq \emptyset$ between the fluid domain $\Omega(t)$ and the solid domain $\Xi(t)$, where $\mathrm{cl}(\cdot)$ represents set closure.

To facilitate the presentation, we denote by $G(A) := \{(A(t), t) : t \in]0, T[\}$ the *space-time* domain corresponding to the time-dependent set $A(t)$.

2.2 Fluid Subsystem

Denoting by r, E, θ and $\boldsymbol{u} = (u_1, \ldots, u_d)$ the density, total energy, absolute temperature and velocity of the fluid, respectively, conservation of mass, momentum and energy is expressed by the Navier-Stokes equations:

$$r' + \partial_j(ru_j) = 0, \qquad (\boldsymbol{x}, t) \in G(\Omega), \tag{1a}$$

$$(ru_i)' + \partial_j(ru_iu_j - \pi_{ij}) = 0, \qquad (\boldsymbol{x}, t) \in G(\Omega), \tag{1b}$$

$$E' + \partial_j(Eu_j - \pi_{ij}u_i - \varkappa\theta_{,j}) = 0, \qquad (\boldsymbol{x}, t) \in G(\Omega), \tag{1c}$$

where we have used the notation $(\cdot)' = \partial/\partial t$ and $\partial_j = \partial/\partial x_j$ and summation on repeated indices is implied, e.g., $\partial_i u_i$ is the divergence of the vector field $\boldsymbol{u}$. Moreover, $\varkappa$ denotes the thermal conductivity, and π_{ij} represent the components of the (Cauchy) stress tensor $\boldsymbol{\pi}$ with respect to the Cartesian basis $\{\boldsymbol{e}_i\}$.

Closure of the system of equations (1) requires suitable constitutive relations which relate $\boldsymbol{\pi}$ and θ to the state variables r, E and $\boldsymbol{u}$. For the purposes in this paper, it is not necessary to specify these constitutive relations. For further details, we refer to, e.g., Refs. [9, 10]. Moreover, in (1) we have assumed the fluid to be compressible. The analysis below extends however with only minor modifications to incompressible fluids.

The system of partial-differential equations (1) must be complemented with suitable auxiliary conditions, viz., initial and boundary conditions. The boundary conditions at the interface $G(\Gamma)$ are treated separately in section 2.4. The auxiliary conditions on the complementary parts of the space-time boundary are not further elaborated.

2.3 Solid Subsystem

The solid is governed by conservation statements identical to (1), but with intrinsically different constitutive relations. However, for solids the equations are customarily transformed to Lagrangian form, i.e., the equations are expressed with respect to a reference configuration. To formulate the solid subproblem, we assume that for all $t \in [0, T]$ the domain $\Xi(t)$ can be characterized by a differentiable one-to-one mapping (diffeomorphism) from the reference configuration $\hat{\Xi}$ according to $\hat{\boldsymbol{x}} \in \hat{\Xi} \mapsto \boldsymbol{\Phi}(\hat{\boldsymbol{x}}, t) \in \Xi(t)$. The map $\hat{\boldsymbol{x}} \mapsto \boldsymbol{\Phi}(\hat{\boldsymbol{x}}, t)$

associates to each point in the reference configuration its position in the actual configuration. The map $\hat{\boldsymbol{x}} \mapsto \boldsymbol{\Phi}(\hat{\boldsymbol{x}}, t) - \hat{\boldsymbol{x}}$ is referred to as the displacement.

Let $\hat{\boldsymbol{x}} = (\hat{x}_1, \ldots, \hat{x}_d)$ denote a Cartesian coordinate system in the reference configuration. The map $\hat{\boldsymbol{x}} \mapsto \boldsymbol{x} = \boldsymbol{\Phi}(\hat{\boldsymbol{x}}, t)$ renders $(\hat{x}_1, \ldots, \hat{x}_d)$ a general coordinate system on $\Xi(t)$. The covariant base vectors associated with this coordinate system are defined by $\boldsymbol{\epsilon}_\alpha = \partial_\alpha \boldsymbol{\Phi}$, where the Greek index serves to indicate that the derivative acts in the reference configuration, i.e., $\partial_\alpha = \partial/\partial \hat{x}_\alpha$. The contravariant base vectors $\boldsymbol{\epsilon}^\alpha$ are defined by $\boldsymbol{\epsilon}^\alpha \cdot \boldsymbol{\epsilon}_\beta = \delta^\alpha_\beta$. The Jacobian of the transformation $\hat{\boldsymbol{x}} \mapsto \boldsymbol{x}$ is given by $J = \det(\epsilon_{\alpha i})$, where $\epsilon_{\alpha i} = \boldsymbol{\epsilon}_\alpha \cdot \boldsymbol{e}_i$ represent the Cartesian components of the covariant base vectors.

Denoting by $\rho(\hat{\boldsymbol{x}}, t)$ the density of the solid in the actual configuration, and by $\hat{\rho}(\hat{\boldsymbol{x}})$ the density in the reference configuration, conservation of mass is expressed by:

$$\rho(\hat{\boldsymbol{x}}, t) J(\hat{\boldsymbol{x}}, t) = \hat{\rho}(\hat{\boldsymbol{x}}), \qquad (\hat{\boldsymbol{x}}, t) \in G(\hat{\Xi}). \tag{2}$$

Using this expression in the equation of motion for the solid, one obtains

$$\hat{\rho}(\hat{\boldsymbol{x}})\, \Phi_i''(\hat{\boldsymbol{x}}, t) - \partial_\alpha S_i^\alpha(\hat{\boldsymbol{x}}, t) = 0, \qquad (\hat{\boldsymbol{x}}, t) \in G(\hat{\Xi}), \tag{3}$$

where S_i^α represent the first-Piola-Kirchoff components of the stress tensor. The first-Piola-Kirchoff components S_i^α and the Cauchy components σ_{ij} of the stress tensor are related by the identity

$$\sigma_{ij} = J^{-1} S_i^\alpha \epsilon_{\alpha j}; \tag{4}$$

see, e.g., Ref. [8]. One may note that (3) describes the Cartesian components of the acceleration of the solid, Φ_i'', in the reference configuration.

In this treatise we restrict ourselves to isothermal solid processes. This implies that the system of equations (2)–(3) can be closed without an energy-conservation equation. Instead, a constitutive relation is required, which relates the stress tensor to the displacement field. For instance, a common constitutive relation for an isotropic solid is $\sigma_{ij} = C_{ijkl}\varepsilon_{kl}$, with $\boldsymbol{\varepsilon}$ the Green-Lagrange strain tensor, viz., $\varepsilon_{\alpha\beta} = \frac{1}{2}(\epsilon_{\alpha i}\epsilon_{\beta i} - \delta_{\alpha\beta})$. For later reference, however, we note that for isothermal processes the solid complies with the energy-conservation equation

$$(\varphi J)'(\hat{\boldsymbol{x}}, t) - \partial_\alpha(\Phi_i' S_i^\alpha)(\hat{\boldsymbol{x}}, t) = 0, \qquad (\hat{\boldsymbol{x}}, t) \in G(\hat{\Xi}), \tag{5}$$

where $\varphi : G(\hat{\Xi}) \to \mathbb{R}$ represents the total internal energy of the solid.

The partial-differential equation (3) must be complemented with suitable initial and boundary conditions. The boundary conditions at the interface are detailed in section 2.4. The auxiliary conditions at the complementary space-time boundary are not further specified.

2.4 Interface Conditions

The fluid and the solid are interconnected by *interface conditions*. With the restriction to isothermal processes, the interface conditions for the fluid-solid system can be separated into kinematic and dynamic conditions.

The kinematic conditions stipulate that the boundaries of the fluid and the solid domains coincide at the interface,

$$\mathrm{cl}(G(\Omega)) \cap \mathrm{cl}(G(\Xi)) = \mathrm{cl}(G(\Gamma)), \tag{6}$$

and, moreover, that the fluid velocity at the interface is identical to the structural velocity at the interface. The latter condition can be expressed as

$$\boldsymbol{u}(\boldsymbol{\Phi}(\hat{\boldsymbol{x}},t),t) = \boldsymbol{\Phi}'(\hat{\boldsymbol{x}},t), \qquad (\hat{\boldsymbol{x}},t) \in G(\hat{\Gamma})\,. \tag{7}$$

For inviscid flows it suffices that the velocities of the fluid and the solid in the direction normal to the interface are identical. Denoting by $\boldsymbol{n}(\boldsymbol{x},t)$ the outward unit normal vector on the boundary $\partial\Omega(t)$ of the fluid domain, and by $\boldsymbol{\lambda}(\hat{\boldsymbol{x}},t)$ the outward unit normal vector on the boundary $\Xi(t)$ of the solid domain in the actual configuration, the kinematic condition (7) thus reduces to

$$\boldsymbol{u}(\boldsymbol{\Phi}(\hat{\boldsymbol{x}},t),t) \cdot \boldsymbol{n}(\boldsymbol{\Phi}(\hat{\boldsymbol{x}},t),t) = -\boldsymbol{\Phi}'(\hat{\boldsymbol{x}},t) \cdot \boldsymbol{\lambda}(\hat{\boldsymbol{x}},t), \qquad (\hat{\boldsymbol{x}},t) \in G(\hat{\Gamma})\,.$$

The dynamic condition specifies the equilibrium of stresses at the interface. This translates into the condition

$$\begin{aligned}\pi_{ij}(\boldsymbol{\Phi}(\hat{\boldsymbol{x}},t),t)\, n_j(\boldsymbol{\Phi}(\hat{\boldsymbol{x}},t),t)\, \mathrm{d}\mu_{\partial\Omega}(\boldsymbol{\Phi}(\hat{\boldsymbol{x}},t),t)& \\ = -S_i^\alpha(\hat{\boldsymbol{x}},t)\, \hat{\lambda}_\alpha(\hat{\boldsymbol{x}},t)\, \mathrm{d}\mu_{\partial\hat{\Xi}}(\hat{\boldsymbol{x}},t), \qquad (\hat{\boldsymbol{x}},t) \in G(\hat{\Gamma})\,,&\end{aligned} \tag{8}$$

with $\hat{\lambda}_\alpha$ the Cartesian components of the unit normal vector on the boundary of the solid domain in the reference configuration. Furthermore, $\mathrm{d}\mu_{\partial\Omega}$ and $\mathrm{d}\mu_{\partial\hat{\Xi}}$ denote the surface measures carried by the boundary of the fluid domain and the boundary of the structural domain in the reference configuration, respectively. It is noteworthy that Nanson's formula,

$$\lambda_i\, \mathrm{d}\mu_{\partial\Xi} = J\, (\partial\hat{x}_j/\partial x_i)\, \hat{\lambda}_j\, \mathrm{d}\mu_{\partial\hat{\Xi}},$$

(see, e.g., [7, p. 88]) and (4) yield the following sequence of identities:

$$S_i^\alpha\, \hat{\lambda}_\alpha\, \mathrm{d}\mu_{\partial\hat{\Xi}} = S_i^\alpha\, \frac{\partial x_j}{\partial\hat{x}_\alpha}\, \frac{\partial\hat{x}_k}{\partial x_j}\, \hat{\lambda}_k\, \mathrm{d}\mu_{\partial\hat{\Xi}} = S_i^\alpha\, \epsilon_{\alpha j}\, J^{-1}\, \lambda_j\, \mathrm{d}\mu_{\partial\Xi} = \sigma_{ij}\, \lambda_j\, \mathrm{d}\mu_{\partial\Xi}. \tag{9}$$

Hence, equation (8) indeed imposes the equilibrium of tractions at the interface, with the fluid contribution expressed in the actual configuration and the solid contribution expressed in the reference configuration. Let us allude to the fact that for inviscid fluids the left member of (8) reduces to pn_i and, accordingly, it holds that $w_i\pi_{ij}n_j = pw_in_i = 0$ for all vectors $\boldsymbol{w} \in T_{\boldsymbol{x}}\Gamma(t)$

tangent to the interface. This implies that the shear stresses vanish at the interface.

In principle, the fluid and solid subsystems also interact through thermodynamic interface conditions. With the restriction to isothermal processes in the solid, however, the initial-boundary-value problem for the fluid must instead be complemented separately with thermodynamic boundary conditions, for instance, a specification of the temperature or the heat flux at the interface.

2.5 Conservation Properties

To establish the conservation properties of the aggregated fluid-solid system, we consider the generic conservation statement $w' + \partial_i f_i = 0$ on the arbitrary space-time domain $G(A)$. Integration on the domain $G(A)$ and application of the divergence theorem yields

$$\left[\int_{A(t)} w(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}\right]_0^T + \int_{G(\partial A)} w\nu + f_i\bar{n}_i\,\mathrm{d}\mu_{G(\partial A)} = 0, \tag{10}$$

with $(\bar{n}_1,\ldots,\bar{n}_d,\nu)$ the space-time normal vector on $G(\partial A)$ and $\mathrm{d}\mu_{G(\partial A)}$ the measure carried by the space-time boundary. On account of the identities

$$\bar{n}_i = n_i(1+v_n^2)^{-1/2},\ \nu = -(1+v_n^2)^{-1/2}v_n,\ \mathrm{d}\mu_{G(\partial A)} = (1+v_n^2)^{1/2}\,\mathrm{d}\mu_{\partial A(t)}\,\mathrm{d}t,$$

with v_n the velocity of the boundary $\partial A(t)$ in its normal direction, Eq. (10) implies that

$$\left[\int_{A(t)} w(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}\right]_0^T = \int_0^T\int_{\partial A(t)} wv_n\,\mathrm{d}\mu_{\partial A(t)}\,\mathrm{d}t - \int_0^T\int_{\partial A(t)} f_i n_i\,\mathrm{d}\mu_{\partial A(t)}\,\mathrm{d}t. \tag{11}$$

By applying Eqs. (10)–(11) to the conservation equations for the fluid and the solid (1)–(5), and combining the resulting expressions according to their connotation, we obtain

$$\left[\int_{\Omega(t)} r(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x} + \int_{\hat{\Xi}} \hat{\rho}(\hat{\boldsymbol{x}})\,\mathrm{d}\hat{\boldsymbol{x}}\right]_0^T = \varrho_1 + \psi_1, \tag{12a}$$

$$\left[\int_{\Omega(t)} [ru_i](\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x} + \int_{\hat{\Xi}} \hat{\rho}(\hat{\boldsymbol{x}})\,\Phi_i'(\hat{\boldsymbol{x}},t)\,\mathrm{d}\hat{\boldsymbol{x}}\right]_0^T = \varrho_{2i} + \psi_{2i}, \tag{12b}$$

$$\left[\int_{\Omega(t)} E(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x} + \int_{\hat{\Xi}} \Phi(\hat{\boldsymbol{x}},t)\,\mathrm{d}\hat{\boldsymbol{x}}\right]_0^T = \varrho_3 + \psi_3, \tag{12c}$$

where the various ψ terms pertain to integrals on the external boundary of the aggregated fluid-solid system and the *interface conservation residuals* are defined as:

$$\varrho_1 = \int_0^T \int_{\Gamma(t)} r(v_n - u_j n_j)\,\mathrm{d}\mu_{\Gamma(t)}\,\mathrm{d}t, \tag{13a}$$

$$\varrho_{2i} = \int_0^T \int_{\Gamma(t)} r u_i (v_n - u_j n_j)\,\mathrm{d}\mu_{\Gamma(t)}\,\mathrm{d}t + \int_0^T \int_{\Gamma(t)} \pi_{ij} n_j\,\mathrm{d}\mu_{\Gamma(t)}\,\mathrm{d}t + \int_0^T \int_{\hat{\Gamma}} S_i^\alpha \hat{\lambda}_\alpha\,\mathrm{d}\mu_{\hat{\Gamma}}\,\mathrm{d}t, \tag{13b}$$

$$\varrho_3 = \int_0^T \int_{\Gamma(t)} E(v_n - u_j n_j)\,\mathrm{d}\mu_{\Gamma(t)}\,\mathrm{d}t + \int_0^T \int_{\Gamma(t)} u_i \pi_{ij} n_j\,\mathrm{d}\mu_{\Gamma(t)}\,\mathrm{d}t + \int_0^T \int_{\hat{\Gamma}} \Phi_i' S_i^\alpha \hat{\lambda}_\alpha\,\mathrm{d}\mu_{\hat{\Gamma}}\,\mathrm{d}t. \tag{13c}$$

One can infer that the residuals in (13) vanish subject to the interface conditions (6)–(8). For example, the integrant in the first term in (13c) vanishes identically on account of the kinematic conditions (6)–(7). Moreover, by transforming the third term in (13c) to the actual solid configuration and adding a suitable partition of zero, we obtain:

$$\begin{aligned}&\int_0^T \int_{\Gamma(t)} u_i \pi_{ij} n_j\,\mathrm{d}\mu_{\Gamma(t)}\,\mathrm{d}t + \int_0^T \int_{\hat{\Gamma}} \Phi_i' S_i^\alpha \hat{\lambda}_\alpha\,\mathrm{d}\mu_{\hat{\Gamma}}\,\mathrm{d}t \\ &= \int_0^T \int_{\Gamma(t)} (\Phi_i' - u_i)\,\sigma_{ij}\lambda_j\,\mathrm{d}\mu_{\Gamma(t)}\,\mathrm{d}t + \int_0^T \int_{\Gamma(t)} (\sigma_{ij}\lambda_j + \pi_{ij} n_j) u_i\,\mathrm{d}\mu_{\Gamma(t)}\,\mathrm{d}t.\end{aligned} \tag{14}$$

It follows from (6)–(9) that the integrants in the right member of (14) vanish identically and, hence, $\varrho_3 = 0$.

In the absence of the residuals, Eqs. (12a)–(12c) can be identified as statements of conservation of mass, momentum and energy for the aggregated fluid-solid system. The interface is absent in the conservation statements for the aggregated system, and changes in the conserved quantities are incurred at the external boundary of the aggregated system only.

2.6 Incompatibility and Conservation

Equipped with the interface residuals (13), we can state the problem considered in this paper more precisely. In a finite-element discretization of a fluid-solid-interaction problem, it is in general prohibitively restrictive to stipulate compatibility of the approximation spaces pertaining to the fluid, the solid and the geometry of the fluid domain. For incompatible approximation spaces, however, the interface conditions cannot be satisfied in a pointwise sense. This implies that for incompatible finite-element approximations, the interface conservation residuals do not generally vanish and, accordingly, the discrete approximation does not maintain the conservation properties of the

continuum problem. The objective of this paper is to devise a methodology for maintaining conservation under incompatibility for finite-element discretizations of fluid-solid-interaction problems.

3 Finite-Element Formulation

3.1 Finite-Element Setting

We restrict ourselves to boundary-fitted finite-element methods. Moreover, we assume that the fluid domain $\Omega(t)$ is always diffeomorphic to the initial domain $\Omega(0)$. Accordingly, $\Omega(t)$ can be characterized by a diffeomorphism $\boldsymbol{\Psi} : \hat{\Omega} \to \Omega(t)$ from the reference domain $\hat{\Omega} = \Omega(0)$. Let $\hat{\boldsymbol{\Omega}}$ represent a finite-element partition of the reference domain into non-overlapping elements e, i.e., $\hat{\boldsymbol{\Omega}}$ is a collection of subdomains D_e, e.g., a simplicial complex [1]. We refer to $|\hat{\boldsymbol{\Omega}}| = \mathrm{int}(\cup_{e\in\hat{\boldsymbol{\Omega}}} \mathrm{cl}(D_e))$ as the underlying space of $\hat{\boldsymbol{\Omega}}$, and we assume that $|\hat{\boldsymbol{\Omega}}| \approx \hat{\Omega}$ in the sense that $|\hat{\boldsymbol{\Omega}}|$ and $\hat{\Omega}$ are homeomorphic and close, e.g., in terms of their Hausdorff distance. Let $V(|\hat{\boldsymbol{\Omega}}|)$ denote a finite-element approximation space for the fluid-domain map $\boldsymbol{\Psi}$. For instance, $V(|\hat{\boldsymbol{\Omega}}|)$ can consist of all continuous functions $|\hat{\boldsymbol{\Omega}}| \mapsto \mathbb{R}^d$ that are piecewise polynomial of a certain degree on the elements D_e, e.g., iso-parametric elements [6]. Similarly, $\hat{\boldsymbol{\Xi}}$ represents a finite-element partition of the solid reference domain $\hat{\Xi}$ and $W(|\hat{\boldsymbol{\Xi}}|)$ is a finite-element space for the solid position.

Let $\partial|\hat{\boldsymbol{\Omega}}|_\Gamma$ and $\partial|\hat{\boldsymbol{\Xi}}|_\Gamma$ represent the sections of the underlying-domain boundaries associated with the interface for the fluid and the solid, respectively. We refer to $\partial|\hat{\boldsymbol{\Omega}}|_\Gamma$ and $\partial|\hat{\boldsymbol{\Xi}}|_\Gamma$ as the interface boundaries. At the interface, the fluid and solid boundaries can assume positions in the trace spaces:

$$V_\Gamma = V_\Gamma(\partial|\hat{\boldsymbol{\Omega}}|_\Gamma) = \{\gamma_{\partial|\hat{\boldsymbol{\Omega}}|_\Gamma}\boldsymbol{\Psi} : \boldsymbol{\Psi} \in V(|\hat{\boldsymbol{\Omega}}|)\},$$
$$W_\Gamma = W_\Gamma(\partial|\hat{\boldsymbol{\Xi}}|_\Gamma) = \{\gamma_{\partial|\hat{\boldsymbol{\Xi}}|_\Gamma}\boldsymbol{\Phi} : \boldsymbol{\Phi} \in W(|\hat{\boldsymbol{\Xi}}|)\},$$

where the trace operators $\gamma_{\partial|\hat{\boldsymbol{\Omega}}|_\Gamma}$ and $\gamma_{\partial|\hat{\boldsymbol{\Xi}}|_\Gamma}$ represent the extension of functions in $V(|\hat{\boldsymbol{\Omega}}|$ and $W(|\hat{\boldsymbol{\Xi}}|)$ onto the fluid and solid interface boundaries, respectively. The tractions that occur at the fluid and solid interface boundaries reside in the dual spaces V_Γ' and W_Γ', respectively. Incompatibility of the finite-element spaces can be characterized as $V_\Gamma \neq W_\Gamma$ and $V_\Gamma' \neq W_\Gamma'$.

In this paper we consider tensor-product space-time finite-element discretizations. Denoting by $V(A)$ a spatial finite-element space on an arbitrary domain A and by $P^\kappa(\tau)$ the space of polynomials of degree $\leq \kappa$ on the generic temporal interval $0 < t < \tau$, the corresponding tensor-product space-time finite-element space is given by $V(A) \times P^\kappa(\tau)$. Let us remark that the analysis below extends mutatis mutandis to conventional finite-element discretizations in which space and time are treated separately.

3.2 Mesh Association

A profound complication in the treatment of incompatible discretizations, is the disparity of the geometric representations of the interface corresponding to the fluid and the solid. The expressions for the interface conditions (7)–(8) are contingent on the fact that the fluid and solid interface boundaries coincide and, accordingly, that $\boldsymbol{\Phi}$ provides a one-to-one association between the solid interface boundary in the reference configuration, $\partial\hat{\Xi}_\Gamma$, and the fluid interface boundary, $\partial\Omega(t)_\Gamma$. Under incompatibility, the fluid and solid interface boundaries do not generally coincide, which necessitates the construction of a proper associative map between the two boundaries. The construction of an association between two geometrically distinct meshes is commonly referred to as *mesh association*. Common examples of mesh-association algorithms are *nearest-neighbor* and *orthogonal-projection* association, although these associations are not generally one-to-one. A detailed treatment of mesh association is beyond the scope of this paper. For further elaboration, we refer to [2, 4]. In the following exposition, a proper associative map $\Lambda : \partial|\hat{\boldsymbol{\Xi}}|_\Gamma \to \partial|\hat{\boldsymbol{\Omega}}|_\Gamma$ is a proviso.

3.3 Boundary Conditions

The interface conditions can be conceived as boundary conditions for the fluid and solid subproblems on the interface boundaries. We adhere to the conventional partitioning of the interface conditions: the kinematic condition (7) yields essential (Dirichlet) boundary conditions for the fluid subsystem and the dynamic condition (8) provides natural (Neumann) boundary conditions for the solid subsystem.

To elucidate the enforcement of natural boundary conditions in the finite-element formulation of the solid, we suppose that the boundary $\partial|\hat{\boldsymbol{\Xi}}|$ is composed of a nonempty section $\partial|\hat{\boldsymbol{\Xi}}|_D$ on which essential boundary conditions are specified, and a section $\partial|\hat{\boldsymbol{\Xi}}|_N \supseteq \partial|\hat{\boldsymbol{\Xi}}|_\Gamma$ on which natural boundary conditions are prescribed. For simplicity, we assume homogeneous boundary conditions on $\partial|\hat{\boldsymbol{\Xi}}|_D$ and $\partial|\hat{\boldsymbol{\Xi}}|_N = \partial|\hat{\boldsymbol{\Xi}}|_\Gamma$, but these assumptions are nonessential. Denoting by $W_0(|\hat{\boldsymbol{\Xi}}|)$ the subspace of $W(|\hat{\boldsymbol{\Xi}}|)$ of functions that vanish on $\partial|\hat{\boldsymbol{\Xi}}|_D$, the space-time finite-element formulation of the solid subsystem can be condensed into the variational formulation: Find $\boldsymbol{\Phi} \in W_0(|\hat{\boldsymbol{\Xi}}|) \times P^\kappa(\tau)$ such that

$$\mathcal{S}(\boldsymbol{\Phi}; \boldsymbol{Y}) = \ell(\boldsymbol{Y}) + \imath(\boldsymbol{Y}) \qquad \forall \boldsymbol{Y} \in W_0(|\hat{\boldsymbol{\Xi}}|) \times P^\kappa(\tau), \tag{15}$$

where

$$\mathcal{S}(\boldsymbol{\Phi}, \boldsymbol{Y}) = \int_0^\tau \int_{|\hat{\boldsymbol{\Xi}}|} \big[Y_i\, \hat{\rho}\, \Phi_i'' + S_i^\alpha\, \partial_\alpha Y_i\big](\hat{\boldsymbol{x}}, t)\, \mathrm{d}\hat{\boldsymbol{x}}\, \mathrm{d}t + \int_{|\hat{\boldsymbol{\Xi}}|} \big[Y_i \Phi_i' - Y_i' \Phi_i\big](\hat{\boldsymbol{x}}, 0)\, \mathrm{d}\hat{\boldsymbol{x}}, \tag{16a}$$

$$\ell(\boldsymbol{Y}) = \int_0^\tau \int_{\partial|\hat{\boldsymbol{\Xi}}|_N} [Y_i \hat{s}_i](\hat{\boldsymbol{x}}, t)\, \mathrm{d}\mu_{\partial|\hat{\boldsymbol{\Xi}}|}(\hat{\boldsymbol{x}})\, \mathrm{d}t, \tag{16b}$$

$$\imath(\boldsymbol{Y}) = \int_{|\hat{\boldsymbol{\Xi}}|} \big(Y_i(\hat{\boldsymbol{x}}, 0)\, \boldsymbol{\Theta}_1(\hat{\boldsymbol{x}}) - Y_i'(\hat{\boldsymbol{x}}, 0)\, \boldsymbol{\Theta}_0(\hat{\boldsymbol{x}}) \big)\, \mathrm{d}\hat{\boldsymbol{x}}. \tag{16c}$$

The function $\hat{\boldsymbol{s}}$ in (16b) represents the traction data in the reference configuration on the boundary $\partial|\hat{\boldsymbol{\Xi}}|_N$ and $\boldsymbol{\Theta}_1$ and $\boldsymbol{\Theta}_0$ in (16c) represent initial data according to $\boldsymbol{\Phi}(\cdot, 0) = \boldsymbol{\Theta}_0$ and $\boldsymbol{\Phi}'(\cdot, 0) = \boldsymbol{\Theta}_1$. In formulation (15)–(16), the initial conditions are enforced weakly. However, other formulations are possible, e.g., with strongly enforced initial conditions. The most important term for our considerations is the *load functional* (16b), which accounts for the weak enforcement of the load data on the interface boundary. In (16b), we conceive of ℓ as a linear functional from $W_0(|\hat{\boldsymbol{\Xi}}|) \times P^\kappa(\tau)$ into $\mathbb{R}$. It is to be noted, however, that $\ell(\boldsymbol{Y})$ depends on the trace of $\boldsymbol{Y}$ at $\partial|\hat{\boldsymbol{\Xi}}|_\Gamma$ only. The following identity holds for all $\boldsymbol{Y} \in W_0(|\hat{\boldsymbol{\Xi}}|) \times P^\kappa(\tau)$:

$$\int_0^\tau \int_{\partial|\hat{\boldsymbol{\Xi}}|_\Gamma} Y_i\, S_i^\alpha \hat{\lambda}_\alpha\, \mathrm{d}\mu_{\partial|\hat{\boldsymbol{\Xi}}|_\Gamma}\, \mathrm{d}t = \ell(\boldsymbol{Y})\,.$$

In particular, the right-most terms in the interface conservation residuals (13b) and (13c) are to be interpreted as

$$\int_0^\tau \int_{\hat{\Gamma}} S_i^\alpha \hat{\lambda}_\alpha\, \mathrm{d}\mu_{\hat{\Gamma}}\, \mathrm{d}t = \ell(\underline{\boldsymbol{e}}_i), \qquad \int_0^\tau \int_{\hat{\Gamma}} \Phi_i' S_i^\alpha \hat{\lambda}_\alpha\, \mathrm{d}\mu_{\hat{\Gamma}}\, \mathrm{d}t = \ell(\underline{\boldsymbol{\Phi}}')\,, \tag{17}$$

where $\underline{\boldsymbol{e}}_i$ and $\underline{\boldsymbol{\Phi}}'$ indicate functions in $W_0(|\hat{\boldsymbol{\Xi}}|) \times P^\kappa(\tau)$ such that $\gamma_{\partial|\hat{\boldsymbol{\Xi}}|_\Gamma} \underline{\boldsymbol{e}}_i$ coincides almost everywhere with any one of the Cartesian base vectors $\boldsymbol{e}_i$ (cf. section 2.1) and $\gamma_{\partial|\hat{\boldsymbol{\Xi}}|_\Gamma} \underline{\boldsymbol{\Phi}}' = \gamma_{\partial|\hat{\boldsymbol{\Xi}}|_\Gamma} \boldsymbol{\Phi}'$ almost everywhere.

We assume that the essential boundary conditions for the fluid are enforced strongly in such a manner that $\gamma_{\partial|\hat{\boldsymbol{\Omega}}|_\Gamma} \boldsymbol{u} = \gamma_{\partial|\hat{\boldsymbol{\Omega}}|_\Gamma} \boldsymbol{\Psi}'$ identically, i.e., the fluid velocity at the fluid interface boundary is identical to the velocity of the fluid interface boundary. This identity stipulates compatibility between the approximation spaces for the fluid-domain position, $\boldsymbol{\Psi}$, and the fluid-velocity, $\boldsymbol{u}$. For instance, this compatibility condition is satisfied if the approximation space for the fluid velocity at the interface is $V_\Gamma \times P^\kappa(\tau)$ and the approximation space for the fluid-domain position at the interface is $V_\Gamma \times P^{\kappa+1}(\tau)$, because $(\cdot)'$ maps $P^{\kappa+1}(\tau)$ into $P^\kappa(\tau)$.

4 Conservation under Incompatibility: the NPCL Method

Our approach to maintaining conservation under incompatibility consists of three complementary components, viz., a suitable nonlinear variational projection that imposes the kinematic condition (7), Riesz' representation of the

load functional in the velocity space, and a consistent evaluation of the load functional (16b) that accounts for the dynamic condition (8). We refer to the approach concisely as the NPCL (Nonlinear variational Projection with Consistent Loading) method.

4.1 Riesz Representation of Load Functionals

An important primitive in the NPCL method is the identification of the traction exerted by the fluid on the solid to an element of the velocity space by means of Riesz' representation theorem. To elaborate the identification, let H denote a Hilbert space with inner product $(\cdot,\cdot)_H$. Riesz' representation theorem asserts that for any linear functional $f \in H'$ there exists a unique element $\mathcal{J}(f) \in H$ such that

$$(\mathcal{J}(f), v)_H = f(v) \qquad \forall v \in H\,,$$

and, moreover, $\|\mathcal{J}(f)\|_H = \|f\|'_H$; see, e.g., [6, 11]. A Hilbert space H can therefore be identified with its dual by means of the ambiguous notation $\mathcal{J}(f) = f$. A Hilbert space H endowed with the identification $H' = H$ is called a pivot space. Moreover, if V is a closed subspace of the pivot space H, then V' can be identified with a superspace of H.

The above expose reveals that the L^2 inner product of displacements and tractions in the load functional (16b) is to be understood as a duality pairing, under the convention that L^2 acts as pivot space. Hence, the load functional admits a Riesz representation in the *displacement* space. Conservation of energy, however, is based on the duality pairing $\langle \ell, \boldsymbol{\Phi}' \rangle$ of the load functional with the interface *velocity* and, remarkably, displacement and velocity do not generally reside in the same function space; cf. [5, §3.4 and §3.8]. More precisely, the displacement space can generally be embedded in the velocity space. This implies that one can only assign meaning to the duality pairing of velocity and load that underlies conservation of energy under additional assumptions on the regularity of the velocities and tractions. Assuming that the load functional in fact resides in the dual velocity space, ℓ admits a Riesz representation according to

$$(\boldsymbol{\pi}_n, \boldsymbol{v})_{\mathscr{V}} = \ell(\boldsymbol{v}) \qquad \forall \boldsymbol{v} \in \mathscr{V}, \tag{18}$$

where $\mathscr{V}$ denotes the velocity function space (on a space-time domain) and $(\cdot,\cdot)_{\mathscr{V}}$ represents the corresponding inner product. Riesz' representation theorem moreover asserts that the identity (18) defines a unique element $\boldsymbol{\pi}_n \in \mathscr{V}$.

4.2 The NPCL Method

In the NPCL method, the kinematic condition (7) is imposed by defining the velocity of the fluid interface boundary as the projection of the velocity of the solid interface boundary. The projection can be condensed into the variational problem: Find $\boldsymbol{u} \in V_\Gamma \times P^\kappa(\tau)$ such that

$$\int_0^\tau \big(\boldsymbol{u}(\cdot,t),\boldsymbol{v}(\cdot,t)\big)_{V,g(t)}\,\mathrm{d}t = \int_0^\tau \big(\boldsymbol{\Phi}'(\Lambda^{-1}(\cdot),t),\boldsymbol{v}(\cdot,t)\big)_{V,g(t)}\,\mathrm{d}t \qquad \forall \boldsymbol{v}\in V_\Gamma\times P^\kappa(\tau). \quad (19)$$

The position of the fluid interface boundary follows by integration. The inner product $(\cdot,\cdot)_{V,g(t)}$ represents the spatial inner product pertaining to the velocity space. This inner product depends on the *metric* g of the actual fluid interface boundary, viz., a twice-covariant vector field that expresses the distance on the boundary. For example, if $(\cdot,\cdot)_{V,g(t)}$ is the L^2 inner product, then

$$\int_0^\tau \big(\boldsymbol{u}(\cdot,t),\boldsymbol{v}(\cdot,t)\big)_{V,g(t)}\,\mathrm{d}t = \int_0^\tau \int_{\partial|\hat{\boldsymbol{\Omega}}|_\Gamma} \boldsymbol{u}(\boldsymbol{x},t)\,\boldsymbol{v}(\boldsymbol{x},t)\,\mathrm{d}\mu_{\partial\Omega(t)}(\boldsymbol{x})\,\mathrm{d}t\,,$$

and $\mathrm{d}\mu_{\partial\Omega(t)}(\boldsymbol{x}) = \sqrt{\det(g)}\,\mathrm{d}\boldsymbol{x}$; see section 5 for exemplification. It is to be noted that the projection (19) is nonlinear, on account of the fact that the metric g depends on the actual position of the fluid interface boundary, which in turn depends on $\boldsymbol{u}$. Moreover, the integrals in (19) are evaluated on the fluid interface boundary, which necessitates the transfer of the velocity on the solid interface boundary to the fluid interface boundary in the right member of (19). This transfer is accomplished by the inverse of the associative map $\Lambda : \partial|\hat{\boldsymbol{\Xi}}|_\Gamma \to \partial|\hat{\boldsymbol{\Omega}}|_\Gamma$; see section 3.2. Note that the association is established in the reference configuration and, hence, Λ is independent of $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$.

In the NPCL method, the dynamic condition (8) is imposed by means of the following specification of the load functional, consistent with (19),

$$\ell(\boldsymbol{Y}) = -\int_0^\tau \big(\boldsymbol{Y}(\Lambda^{-1}(\cdot),t),\boldsymbol{\pi}_n(\cdot,t)\big)_{V,g(t)}\,\mathrm{d}t \quad (20)$$

where $\boldsymbol{\pi}_n$ designates the approximation of the Riesz representation of the traction $\pi_{ij}n_j$ in the velocity space, defined as the unique element of $V_\Gamma\times P^\kappa(\tau)$ in compliance with

$$\int_0^\tau \big(\boldsymbol{\pi}_n(\cdot,t),\boldsymbol{v}(\cdot,t)\big)_{V,g(t)}\,\mathrm{d}t = \int_0^\tau \int_{\partial|\hat{\boldsymbol{\Omega}}|_\Gamma} [\pi_{ij}n_j]\,v_i(\boldsymbol{x},t)\;\mathrm{d}\mu_{\partial\Omega(t)}(\boldsymbol{x})\,\mathrm{d}t \qquad \forall \boldsymbol{v}\in V_\Gamma\times P^\kappa(\tau)\,. \quad (21)$$

The functional ℓ according to (20) depends implicitly on the position of the fluid interface boundary through the metric g. The evaluation of the test functions pertaining to the solid, $\boldsymbol{Y}$, on the fluid interface boundary again requires the inverse association Λ^{-1}.

To prove that (19)–(21) ensure conservation of mass, momentum and energy, we first note that the first terms in the interface conservation residuals (13) vanish on account of the stipulated identity $\gamma_{\partial|\hat{\boldsymbol{\Omega}}|_\Gamma}\boldsymbol{u} = \gamma_{\partial|\hat{\boldsymbol{\Omega}}|_\Gamma}\boldsymbol{\Psi}'$. Hence, $\varrho_1 = 0$ and conservation of mass is ensured. Moreover, the first identity in (17) and equations (20) and (21) yield the sequence of identities:

$$\int_0^\tau \int_{\hat{\Gamma}} S_i^\alpha \hat{\lambda}_\alpha \,\mathrm{d}\mu_{\hat{\Gamma}}\,\mathrm{d}t = \ell(\underline{\boldsymbol{e}}_i) = -\int_0^\tau \big(\underline{\boldsymbol{e}}_i(\Lambda^{-1}(\cdot),t), \boldsymbol{\pi}_n(\cdot,t)\big)_{V,g(t)}\,\mathrm{d}t$$
$$= -\int_0^\tau \int_{\partial|\hat{\boldsymbol{\Omega}}|_\Gamma} [\pi_{ij}n_j](\boldsymbol{x},t)\;\mathrm{d}\mu_{\partial\Omega(t)}(\boldsymbol{x})\,\mathrm{d}t\,. \quad (22)$$

Noting that $\Gamma(t)$ in (13) in fact represents the interface boundary of the fluid domain, equation (22) implies that the second and third term in (13b) cancel. Therefore, ϱ_{2i} vanishes, which implies that momentum is conserved. It is to be mentioned that the second identity in (22) is found on the assumption that $V_\Gamma \times P^\kappa(\tau)$ contains functions that coincide almost everywhere with the Cartesian base vector $\boldsymbol{e}_i$, but this assumption is nonrestrictive. Finally, on account of, successively, (17), Eq. (20), Eq. (19) and $\boldsymbol{\pi}_n \in V_\Gamma \times P^\kappa(\tau)$, and Eq. (21) and $\boldsymbol{u} \in V_\Gamma \times P^\kappa(\tau)$, it holds that:

$$\int_0^\tau \int_{\hat{\Gamma}} \Phi_i' S_i^\alpha \hat{\lambda}_\alpha \,\mathrm{d}\mu_{\hat{\Gamma}}\,\mathrm{d}t = \ell(\underline{\boldsymbol{\Phi}}')$$
$$= -\int_0^\tau \big(\boldsymbol{\Phi}'(\Lambda^{-1}(\cdot),t), \boldsymbol{\pi}_n(\cdot,t)\big)_{V,g(t)}\,\mathrm{d}t = -\int_0^\tau \big(\boldsymbol{u}(\cdot,t), \boldsymbol{\pi}_n(\cdot,t)\big)_{V,g(t)}\,\mathrm{d}t$$
$$= \int_0^\tau \int_{\partial|\hat{\boldsymbol{\Omega}}|_\Gamma} u_i(\boldsymbol{x},t)\,[\pi_{ij}n_j]\,(\boldsymbol{x},t)\,\mathrm{d}\mu_{\partial\Omega(t)}(\boldsymbol{x})\,\mathrm{d}t\,.$$

Hence, the second and third term in (13c) cancel and $\varrho_3 = 0$, which corroborates that the NPCL method also preserves energy.

In a practical implementation, the integrals in (19)–(21) are typically evaluated by means of Gauss quadrature. The evaluation of the test functions of the solid on the fluid interface boundary in (20) then only requires the association of *quadrature points* on the fluid interface boundary in the reference configuration to corresponding points on the solid interface boundary in the reference configuration. Hence, the NPCL method only requires a point-to-point association algorithm.

The $(\cdot,\cdot)_{V,g(t)}$ inner product in (19)–(21) in general assumes a profoundly difficult form, on account of the fact that it pertains to a trace space. A typical functional setting for the solid problem is:

$$\boldsymbol{\Phi} \in L^2\big(0,\tau; H_0^1(\hat{\Xi})\big), \quad \boldsymbol{\Phi}' \in L^2\big(0,\tau; L^2(\hat{\Xi})\big), \quad \boldsymbol{\Phi}'' \in L^2\big(0,\tau; H^{-1}(\hat{\Xi})\big); \quad (23)$$

see [5, §3.8]. However, according to the trace theorem (see, e.g., [6, Thm. 4.9]), this implies that the appropriate corresponding inner product for the interface-boundary velocity is

$$(\boldsymbol{u},\boldsymbol{v})_{\mathscr{V}} = \int_0^\tau \big(\boldsymbol{u}(\cdot,t), \boldsymbol{v}(\cdot,t)\big)_{H^{-1/2},g(t)}\,\mathrm{d}t\,.$$

Although such an inner product is computable, it is utterly impractical. In a practical implementation, one can therefore opt to replace $(\cdot,\cdot)_{V,g(t)}$ by the

standard L^2 inner product, in violation of the proper functional setting. Such a violation of the functional setting can lead to non-convergence if the trace value $\gamma_{\hat{\Gamma}}\boldsymbol{\Phi}'$ does not reside in $L^2(0,\tau;L^2(\hat{\Gamma}))$, i.e., if $\boldsymbol{\Phi}'$ is highly irregular at the interface boundary. The practical relevance of such extreme cases is however limited. One can infer that the conservation properties of the NPCL are preserved if the inner products in (19)–(21) are consistently replaced by the L^2 inner product.

5 Example

To exemplify the NPCL method, we consider its application to an elementary projection problem in (2+1)D space-time. Let us consider approximate fluid and solid interface boundaries in the reference configuration, given by:

$$\partial|\hat{\boldsymbol{\Xi}}|_\Gamma = \left\{\boldsymbol{x}\in\mathbb{R}^2 : \boldsymbol{x} = (s,0),\ 0\le s\le 1\right\},$$
$$\partial|\hat{\boldsymbol{\Omega}}|_\Gamma = \left\{\boldsymbol{x}\in\mathbb{R}^2 : \boldsymbol{x} = \left(s-\tfrac{1}{10},\tfrac{1}{10}s+\tfrac{1}{10}\right),\ 0\le s\le 1\right\}.$$

Clearly, there is a mismatch between the approximate interface boundaries. We construct the following associative map (with corresponding inverse) between $\partial|\hat{\boldsymbol{\Xi}}|_\Gamma$ and $\partial|\hat{\boldsymbol{\Omega}}|_\Gamma$:

$$\Lambda\boldsymbol{x} = \left(x_1-\tfrac{1}{10},x_2+\tfrac{1}{10}x_1+\tfrac{1}{10}\right), \tag{24a}$$
$$\Lambda^{-1}\boldsymbol{x} = \left(x_1+\tfrac{1}{10},x_2-\tfrac{1}{10}x_1-\tfrac{11}{100}\right). \tag{24b}$$

One easily infers that Λ according to (24a) indeed provides a diffeomorphism from $\partial|\hat{\boldsymbol{\Xi}}|_\Gamma$ onto $\partial|\hat{\boldsymbol{\Omega}}|_\Gamma$. In addition, suppose that the position of the solid interface boundary is given by:

$$\boldsymbol{\Phi}(s,t) = (s,0) + \left(\tfrac{1}{10}\sin(2\pi t)\sin(s), t\sin(2\pi s)\right). \tag{25}$$

It is to be noted that due to the initial mismatch between the fluid and solid interface boundaries, the corresponding actual position of the fluid interface boundary is:

$$\boldsymbol{\Psi}(s,t) = (s-\tfrac{1}{10},\tfrac{1}{10}s+\tfrac{1}{10}) + \left(\tfrac{1}{10}\sin(2\pi t)\sin(s), t\sin(2\pi s)\right); \tag{26}$$

see figure 2 on page 17.

We approximate the velocity field corresponding to (25) on the fluid interface boundary $\partial|\hat{\boldsymbol{\Omega}}|_\Gamma$ by tensor product polynomials of degree κ. In particular,

$$\tilde{\boldsymbol{u}}(s,t) = \sum_{i=0}^{\kappa}\sum_{j=0}^{\kappa}\hat{\boldsymbol{u}}_{ij}\eta_i(s)\eta_j(t), \tag{27}$$

where η_i represents the Legendre polynomial of degree i on the interval $[0,1]$. For given coefficients $\hat{\boldsymbol{u}}_{ij}$, the corresponding approximation to the position of the fluid interface boundary is determined by

$$\tilde{\boldsymbol{\Psi}}(s,t) = \left(s - \tfrac{1}{10}, \tfrac{1}{10}s + \tfrac{1}{10}\right) + \sum_{i=0}^{\kappa}\sum_{j=0}^{\kappa} \hat{\boldsymbol{u}}_{ij}\eta_i(s)H_j(t)\,, \tag{28}$$

where H_j represents the anti-derivative of η_j. The metric tensor $g(t)$ and the measure $\mathrm{d}\mu_{\partial\Omega(t)}$ associated to (28) are given by

$$g(t)(s) = \frac{\partial\tilde{\boldsymbol{\Psi}}}{\partial s}\cdot\frac{\partial\tilde{\boldsymbol{\Psi}}}{\partial s} = \left|\frac{\partial\tilde{\boldsymbol{\Psi}}}{\partial s}\right|^2, \qquad \mathrm{d}\mu_{\partial\Omega(t)} = \sqrt{g(t)(s)}\,\mathrm{d}s\,, \tag{29}$$

respectively. Setting $(\cdot,\cdot)_V$ in Eqs. (19)–(21) to the L^2 inner product and denoting by $[P^\kappa]^2$ the tensor-product approximation space of order κ on the unit square, the projection problem (19) can be condensed into the following variational problem: find $\tilde{\boldsymbol{u}} \in [P^\kappa]^2$ such that

$$\mathcal{A}(\tilde{\boldsymbol{u}},\tilde{\boldsymbol{u}},\tilde{\boldsymbol{v}}) = \mathcal{B}(\tilde{\boldsymbol{u}},\tilde{\boldsymbol{v}}) \quad \forall\tilde{\boldsymbol{v}} \in [P^\kappa]^2\,, \tag{30}$$

where the trilinear from $\mathcal{A}$ and the bilinear form $\mathcal{B}$ are defined by

$$\mathcal{A}(\tilde{\boldsymbol{u}},\tilde{\boldsymbol{w}},\tilde{\boldsymbol{v}}) = \int_{[0,1]^2} \tilde{\boldsymbol{v}}\cdot\tilde{\boldsymbol{u}}\left|\frac{\partial\tilde{\boldsymbol{\Psi}}(\tilde{\boldsymbol{w}})}{\partial s}\right|\mathrm{d}s\,\mathrm{d}t\,, \tag{31a}$$

$$\mathcal{B}(\tilde{\boldsymbol{w}},\tilde{\boldsymbol{v}}) = \int_{[0,1]^2} \tilde{\boldsymbol{v}}\cdot\boldsymbol{\Phi}\left|\frac{\partial\tilde{\boldsymbol{\Psi}}(\tilde{\boldsymbol{w}})}{\partial s}\right|\mathrm{d}s\,\mathrm{d}t\,, \tag{31b}$$

and the notation $\tilde{\boldsymbol{\Psi}}(\tilde{\boldsymbol{u}})$ has been used to expose the dependence of $\tilde{\boldsymbol{\Psi}}$ on $\tilde{\boldsymbol{u}}$ through (28). It is to be noted that this dependence renders (30) nonlinear. For completeness, we mention that the integrals in this section, e.g., in (31), are approximated by gauss quadrature with $(\kappa+5)^2$ integration points, which renders the quadrature error negligible.

We solve the nonlinear problem (30) by means of successive approximation: For a given initial estimate $\tilde{\boldsymbol{u}}^0 \in [P^\kappa]^2$, for $n = 1, 2, \ldots$ find $\tilde{\boldsymbol{u}}^n \in [P^\kappa]^2$ such that

$$\mathcal{A}(\tilde{\boldsymbol{u}}^n,\tilde{\boldsymbol{u}}^{n-1},\tilde{\boldsymbol{v}}) = \mathcal{B}(\tilde{\boldsymbol{u}}^{n-1},\tilde{\boldsymbol{v}}) \quad \forall\tilde{\boldsymbol{v}} \in [P^\kappa]^2\,. \tag{32}$$

To illustrate the convergence behavior of the successive-approximation method, figure 1 on the next page plots the norm of the residual,

$$\|r_n\| = \sup_{\tilde{\boldsymbol{v}}\in[P^\kappa]^2} \frac{\mathcal{A}(\tilde{\boldsymbol{u}}^n,\tilde{\boldsymbol{u}}^n,\tilde{\boldsymbol{v}}) - \mathcal{B}(\tilde{\boldsymbol{u}}^n,\tilde{\boldsymbol{v}})}{\mathcal{A}(\tilde{\boldsymbol{v}},\tilde{\boldsymbol{u}}^n,\tilde{\boldsymbol{v}})}, \tag{33}$$

versus the iteration counter n for various orders of approximation κ. Figure 1 confirms that the successive approximation method converges. Moreover, the figure shows that the convergence behavior of the method improves as κ increases. To elaborate this property of the method, let us note that:

$$\begin{aligned}\mathcal{A}(\tilde{\boldsymbol{u}}^n,\tilde{\boldsymbol{u}}^n,\tilde{\boldsymbol{v}}) - \mathcal{B}(\tilde{\boldsymbol{u}}^n,\tilde{\boldsymbol{v}}) &= \mathcal{A}(\tilde{\boldsymbol{u}}^n,\tilde{\boldsymbol{u}}^{n-1},\tilde{\boldsymbol{z}}) - \mathcal{B}(\tilde{\boldsymbol{u}}^{n-1},\tilde{\boldsymbol{z}}) \\ &= \mathcal{A}(\tilde{\boldsymbol{u}}^n,\tilde{\boldsymbol{u}}^{n-1},\tilde{\boldsymbol{z}}-\tilde{\boldsymbol{w}}) - \mathcal{B}(\tilde{\boldsymbol{u}}^{n-1},\tilde{\boldsymbol{z}}-\tilde{\boldsymbol{w}})\,,\end{aligned} \tag{34}$$

for all $\tilde{\boldsymbol{w}} \in [P^\kappa]^2$, where

$$\tilde{\boldsymbol{z}} = \tilde{\boldsymbol{v}} \left| \frac{\partial \tilde{\boldsymbol{\Psi}}(\tilde{\boldsymbol{u}}^n)}{\partial s} \right| \left| \frac{\partial \tilde{\boldsymbol{\Psi}}(\tilde{\boldsymbol{u}}^{n-1})}{\partial s} \right|^{-1} .$$

The final identity in (34) is an immediate consequence of (32). Selecting $\tilde{\boldsymbol{w}}$ in (34) as the projection of $\tilde{\boldsymbol{z}}$ onto $[P^\kappa]^2$, it follows that the residual is increasingly effectively reduced by the successive-approximation method as κ increases. Let us note that this argument is general, and not restricted to the considered model problem. Hence, the convergence behavior of the successive-approximation method for the nonlinear variational problem (19) will in general improve if the mesh is refined or the order of approximation is increased, independent of the specifics of the underlying fluid-solid-interaction problem.

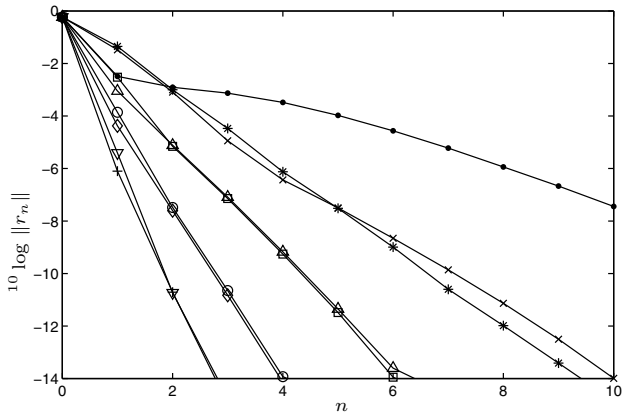


Fig. 1: Convergence behavior of the successive approximation method: norm of the residual according to (33) versus the iteration counter n for $\kappa = 2$ ($\bullet$), $\kappa = 3$ ($*$), $\kappa = 4$ ($\times$), $\kappa = 5$ ($\square$), $\kappa = 6$ ($\triangle$), $\kappa = 7$ ($\circ$), $\kappa = 8$ ($\diamond$), $\kappa = 9$ ($\triangledown$), and $\kappa = 10$ ($+$),

Let us next consider a traction field on the fluid interface boundary with components:

$$\pi_{1j} n_j = 1 + e^t \cos(2\pi s), \qquad \pi_{2j} n_j = 10 + \sin(2\pi s t)\,; \tag{35}$$

see figure 3 on the facing page. The next operation in the NPCL method is the construction of an approximation to the Riesz representation of $\pi_{ij} n_j$ in the velocity space. This operation can be condensed into the variational problem: find $\tilde{\boldsymbol{\pi}} \in [P^\kappa]^2$ such that:

$$\mathcal{A}(\tilde{\boldsymbol{\pi}}, \tilde{\boldsymbol{u}}, \tilde{\boldsymbol{v}}) = \mathcal{L}(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{v}}) \quad \forall \tilde{\boldsymbol{v}} \in [P^\kappa]^2\,, \tag{36}$$

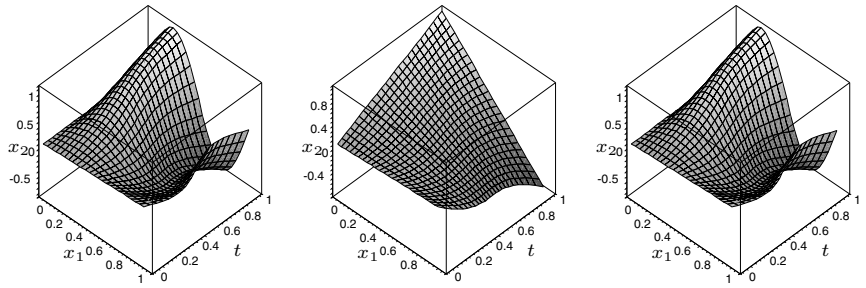


Fig. 2: Position of the fluid interface boundary: exact according to (26) (*left*) and approximations for $\kappa = 2$ (*center*) and for $\kappa = 6$ (*right*).

with the trilinear form $\mathcal{A}$ according to (31a) and the bilinear form $\mathcal{L}$,

$$\mathcal{L}(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{v}}) = \int_{[0,1]^2} \tilde{v}_i \, [\pi_{ij} n_j] \left| \frac{\partial \tilde{\boldsymbol{\Psi}}(\tilde{\boldsymbol{u}})}{\partial s} \right| \mathrm{d}s \, \mathrm{d}t \,. \tag{37}$$

Note that (37) depends on the velocity approximation $\tilde{\boldsymbol{u}}$ defined by (30) via $\tilde{\boldsymbol{\Psi}}(\tilde{\boldsymbol{u}})$.

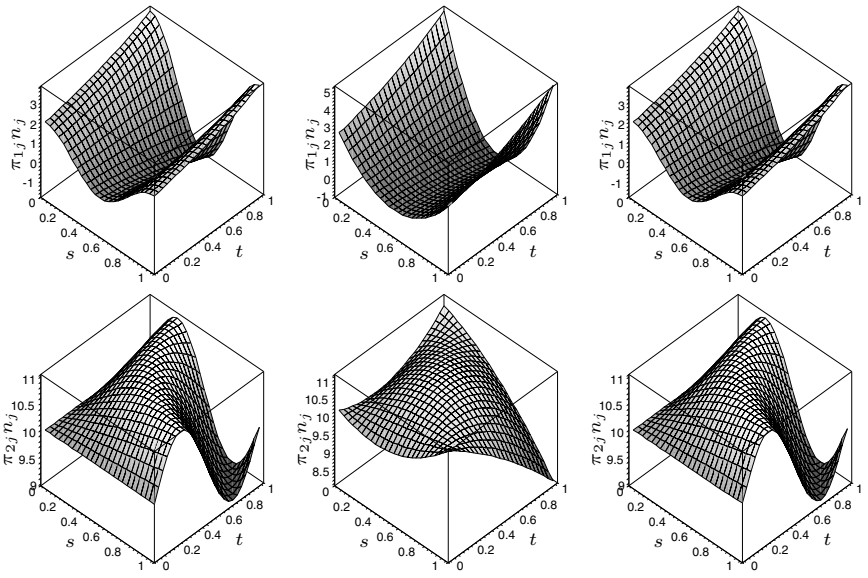


Fig. 3: Traction on the fluid interface boundary, components 1 (*top*) and 2 (*bottom*): exact according to (35) (*left*) and approximations for $\kappa = 2$ (*center*) and for $\kappa = 6$ (*right*).

The final operation in the NPCL method pertains to the construction of the load-functional approximation:

$$\tilde{\ell}(\boldsymbol{Y}) = -\int_{[0,1]^2} \boldsymbol{Y} \cdot \tilde{\boldsymbol{\pi}} \left| \frac{\partial \tilde{\boldsymbol{\Psi}}(\tilde{\boldsymbol{u}})}{\partial s} \right| \mathrm{d}s \, \mathrm{d}t \,. \tag{38}$$

The association between the fluid interface boundary and the solid interface boundary is accounted for by the use of (s,t) coordinates.

To illustrate the conservation properties of the NPCL method, table 1 on the next page lists the energy flux across the fluid interface boundary, $\mathcal{L}(\tilde{\boldsymbol{u}},\tilde{\boldsymbol{u}})$, the momentum flux across the fluid interface boundary, $(\mathcal{L}(\tilde{\boldsymbol{u}},\boldsymbol{e}_1),\mathcal{L}(\tilde{\boldsymbol{u}},\boldsymbol{e}_2))$, the energy flux across the solid interface boundary, $\tilde{\ell}(\boldsymbol{\Phi}')$, and the momentum flux across the solid interface boundary, $(\tilde{\ell}(\boldsymbol{e}_1),\tilde{\ell}(\boldsymbol{e}_2))$, for $\kappa = 2,4,\ldots,10$. For conciseness, the results in table 1 have been restricted to even orders of approximation. Moreover, for $\boldsymbol{\Phi}$, $\boldsymbol{\Psi}$ and $\pi_{ij}n_j$ according to (25), (26) and (35), respectively, we obtain the reference values:

$$\int_{[0,1]^2} \Phi_i' \left[\pi_{ij} n_j\right] \left| \frac{\partial \boldsymbol{\Psi}}{\partial s} \right| \mathrm{d}s \, \mathrm{d}t = \quad 3.606070414460000E-01 \,, \tag{39a}$$

$$\int_{[0,1]^2} \begin{pmatrix} \pi_{1j} n_j \\ \pi_{2j} n_j \end{pmatrix} \left| \frac{\partial \boldsymbol{\Psi}}{\partial s} \right| \mathrm{d}s \, \mathrm{d}t = \begin{pmatrix} 2.481752533540000E+00 \\ 2.464690628610000E+01 \end{pmatrix} . \tag{39b}$$

The results in table 1 corroborate the conservation properties of the NPCL method. For all orders of approximation, the energy and momentum fluxes across the fluid and solid interface boundaries are essentially identical. Moreover, as the order of approximation increases, the approximate fluxes converge to the reference values in (39).

6 Conclusion

We presented a new coupling method for fluid-solid-interaction problems which preserves the conservation properties of the underlying continuum problem under incompatibility of the approximation spaces for the fluid and the solid. The method imposes the kinematic interface condition by means of a nonlinear variational projection of the velocity from the solid boundary to the fluid boundary. The dynamic condition is enforced by a consistent evaluation of the load functional, which involves the Riesz representation of the load functional in the velocity trace space of the fluid. The method is referred to as the NPCL (Nonlinear variational Projection with Consistent Loading) method.

As the appropriate inner product for the velocity trace space can assume a profoundly complicated form, we propounded the use of L^2 inner products in practical computations. The NPCL method preserves its conservation properties if the inner product is consistently replaced in all components of the method.

Table 1: Conservation properties of the NPCL method: energy and momentum fluxes across the fluid and solid interface boundaries for orders of approximation $\kappa = 2, 4, 6, 8, 10$.

	fluid		solid	
$\kappa = 2$	$\mathcal{L}(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{u}})$	$1.538313684748479E-01$	$\tilde{\ell}(\boldsymbol{\Phi}')$	$1.538313684748228E-01$
	$\mathcal{L}(\tilde{\boldsymbol{u}}, \boldsymbol{e}_1)$	$1.395877705335224E+00$	$\tilde{\ell}(\boldsymbol{e}_1)$	$1.395877705335224E+00$
	$\mathcal{L}(\tilde{\boldsymbol{u}}, \boldsymbol{e}_2)$	$1.445000844112377E+01$	$\tilde{\ell}(\boldsymbol{e}_2)$	$1.445000844112375E+01$
$\kappa = 4$	$\mathcal{L}(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{u}})$	$3.830462789356112E-01$	$\tilde{\ell}(\boldsymbol{\Phi}')$	$3.830462789356541E-01$
	$\mathcal{L}(\tilde{\boldsymbol{u}}, \boldsymbol{e}_1)$	$3.199991700241049E+00$	$\tilde{\ell}(\boldsymbol{e}_1)$	$3.199991700241051E+00$
	$\mathcal{L}(\tilde{\boldsymbol{u}}, \boldsymbol{e}_2)$	$2.609101263224458E+01$	$\tilde{\ell}(\boldsymbol{e}_2)$	$2.609101263224456E+01$
$\kappa = 6$	$\mathcal{L}(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{u}})$	$3.763395670325019E-01$	$\tilde{\ell}(\boldsymbol{\Phi}')$	$3.763395670324617E-01$
	$\mathcal{L}(\tilde{\boldsymbol{u}}, \boldsymbol{e}_1)$	$2.463939376383578E+00$	$\tilde{\ell}(\boldsymbol{e}_1)$	$2.463939376383568E+00$
	$\mathcal{L}(\tilde{\boldsymbol{u}}, \boldsymbol{e}_2)$	$2.470803076025526E+01$	$\tilde{\ell}(\boldsymbol{e}_2)$	$2.470803076025531E+01$
$\kappa = 8$	$\mathcal{L}(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{u}})$	$3.547985610065834E-01$	$\tilde{\ell}(\boldsymbol{\Phi}')$	$3.547985610066048E-01$
	$\mathcal{L}(\tilde{\boldsymbol{u}}, \boldsymbol{e}_1)$	$2.473739029165391E+00$	$\tilde{\ell}(\boldsymbol{e}_1)$	$2.473739029165391E+00$
	$\mathcal{L}(\tilde{\boldsymbol{u}}, \boldsymbol{e}_2)$	$2.460630254948687E+01$	$\tilde{\ell}(\boldsymbol{e}_2)$	$2.460630254948681E+01$
$\kappa = 10$	$\mathcal{L}(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{u}})$	$3.598488191723714E-01$	$\tilde{\ell}(\boldsymbol{\Phi}')$	$3.598488191723657E-01$
	$\mathcal{L}(\tilde{\boldsymbol{u}}, \boldsymbol{e}_1)$	$2.483470257439778E+00$	$\tilde{\ell}(\boldsymbol{e}_1)$	$2.483470257439793E+00$
	$\mathcal{L}(\tilde{\boldsymbol{u}}, \boldsymbol{e}_2)$	$2.462978310231236E+01$	$\tilde{\ell}(\boldsymbol{e}_2)$	$2.462978310231259E+01$

Finally, we exemplified the NPCL method by means of numerical experiments on a projection problem. The nonlinear-variational-projection problem was solved by means of successive approximation. We observed that the convergence behavior of the successive-approximation method improves as the approximation space is refined, and we elaborated that this is a generic property of successive approximation for the nonlinear-variational-projection problem, independent of the underlying fluid-solid-interaction problem. This implies that the nonlinear-variational-projection problem in the NPCL method can in general be solved effectively by means of successive approximation. Consideration of the energy and momentum fluxes across the fluid and solid interface boundaries confirmed the conservation properties of the NPCL method.

References

1. E.D. Bloch. *A First Course in Geometric Topology and Differential Geometry.* Birkhäuser, Boston, 1996.
2. E.H. van Brummelen. Mesh association by projection along smoothed-normal-vector fields: association of closed manifolds. 2006. Accepted for publication in Int. J. Num. Meth. Engng. Also available as DACS report DACS-06-005 at http://www.emserver.lr.tudelft.nl/downloads/DACS-06-005.pdf.

3. C. Farhat, M. Lesoinne, and P. LeTallec. Load and motion transfer algorithms for fluid/structure interaction problems with non-matching discrete interfaces: Momentum and energy conservation, optimal discretization and application to aeroelasticity. *Comput. Methods Appl. Mech. Engrg.*, pages 95–114, 1998.
4. X. Jiao and M.T. Heath. Common-refinement-based data transfer between non-matching meshes in multiphysics simulations. *Int. J. Numer. Meth. Engng.*, 61:2402–2427, 2004.
5. J.L. Lions and E. Magenes. *Non-Homogeous Boundary Value Problems and Applications I.* Springer-Verlag, 1972.
6. J.T. Oden and J.N. Reddy. *An Introduction to the Mathematical Theory of Finite Elements.* Pure and Applied Mathematics. John Wiley & Sons, New York, 1974.
7. R.W. Ogden. *Non-linear elastic deformations.* Ellis Horwood, 1984.
8. J. Salençon. *Handbook of Continuum Mechanics.* Springer-Verlag, Heidelberg, 2001.
9. R. Temam and A. Mirainville. *Mathematical Modeling in Continuum Mechanics.* Cambridge University Press, 2001.
10. P. Wesseling. *Principles of Computational Fluid Dynamics*, volume 29 of *Springer Series in Computational Mathematics.* Springer, Berlin, 2001.
11. E. Zeidler. *Applied Functional Analysis: Applications to Mathematical Physics*, volume 108 of *Applied Mathematical Sciences.* Springer, Berlin, 1995.

Author Index

Subject Index

IUTAM Bookseries

1. P. Eberhard (ed.): *IUTAM Symposium on Multiscale Problems in Multibody System Contacts*. 2007 ISBN 978-1-4020-5980-3
2. H. Hu and E. Kreuzer (eds): *IUTAM Symposium on Dynamics and Control of Nonlinear Systems with Uncertainty*. 2007 ISBN 978-1-4020-6331-2
3. P. Wriggers and U. Nackenhorst (eds): *IUTAM Symposium on Computational Methods in Contact Mechanics*. 2007 ISBN 978-4020-6404-3
4. Y. Kaneda (ed.): *IUTAM Symposium on Computational Physics and New Perspectives in Turbulence*. 2007 ISBN 978-4020-6471-5
5. A. Combescure, R. de Borst and T. Belytschko (eds): *IUTAM Symposium on Discretization Methods for Evolving Discontinuities*. 2007 ISBN 978-4020-6529-3