

Pier Francesco Bortignon · Giuseppe Lodato
Emanuela Meroni · Matteo G. A. Paris
Laura Perini · Alessandro Vicini *Editors*

Toward a Science Campus in Milan

A Snapshot of Current Research at the
Physics Department Aldo Pontremoli

 Springer

Toward a Science Campus in Milan

Pier Francesco Bortignon
Giuseppe Lodato · Emanuela Meroni
Matteo G. A. Paris · Laura Perini
Alessandro Vicini
Editors

Toward a Science Campus in Milan

A Snapshot of Current Research at the Physics
Department Aldo Pontremoli

 Springer

Editors

Pier Francesco Bortignon (Deceased)
Milan, Italy

Giuseppe Lodato
Dipartimento di Fisica “Aldo Pontremoli”
Università degli Studi di Milano
Milan, Italy

Emanuela Meroni
Dipartimento di Fisica “Aldo Pontremoli”
Università degli Studi di Milano
Milan, Italy

Matteo G. A. Paris
Dipartimento di Fisica “Aldo Pontremoli”
Università degli Studi di Milano
Milan, Italy

Laura Perini
Dipartimento di Fisica “Aldo Pontremoli”
Università degli Studi di Milano
Milan, Italy

Alessandro Vicini
Dipartimento di Fisica “Aldo Pontremoli”
Università degli Studi di Milano
Milan, Italy

ISBN 978-3-030-01628-9

ISBN 978-3-030-01629-6 (eBook)

<https://doi.org/10.1007/978-3-030-01629-6>

Library of Congress Control Number: 2018958924

© Springer Nature Switzerland AG 2018

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

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

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

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



This volume is dedicated to the memory of our colleague and friend Pier Francesco Bortignon (5/8/1948–27/8/2018), who contributed so much to the success of this workshop with his usual generosity and passion for science.

Preface

A Physics Department is a complex system, more akin to a living organism than a network of weakly interacting units. This is perhaps the reason why being the director of the Physics Department of the University of Milan for about six years has been such an exciting adventure.

I inherited a healthy department from my predecessor, Francesco Ragusa. Since then, we have been growing steadily. Looking at quantitative figures like the number of faculty members, staff members, and students, however, one sees only a side of the story. What I mostly witnessed in the last years is an impressive qualitative growth in terms of the synergy between the different areas of research, the capacity of attracting funds from diverse sources, and the overall throughput of the department's ecosystem.

I deliver to my successor, Giovanni Onida, a department which is just as healthy as before and with a more outward-oriented attitude. This will be a key feature in the next future, which will see the University of Milan on the verge of a radically new era, as the scientific Departments will move to the new Campus at the EXPO2015 site. The Physics Department, recently named after its founder Aldo Pontremoli, has embraced this new challenge as an opportunity to expand its activities, possibly including the development of new interdisciplinary facilities for applied physics.

With this idea very much in mind, I have been delighted to conclude my service as a director with a Congress of the Department. I'm also proud to present this book of proceeding, which contains a faithful account of what we have seen and heard during our Congress.

I would like to thank the Scientific Committee, all the speakers, and the many participants, for ensuring the high scientific quality of the workshop and of this volume. I warmly thank also the colleagues at the Mathematics Department, for having hosted our workshop in their conference hall.

Finally, special thanks are due to the local organizers Vera Bernardoni and Matteo Bina, for their help and patience in preparing and running the workshop.

Milan, Italy
March 2018

Laura Perini

Contents

1	Measuring the Universe with Galaxy Redshift Surveys	1
	L. Guzzo, J. Bel, D. Bianchi, C. Carbone, B. R. Granett, A. J. Hawken, F. G. Mohammad, A. Pezzotta, S. Rota and M. Zennaro	
2	Theoretical Spectroscopy Characterization of Deep Electronic States of Defects in Silicon via Density Functional Theory with Hybrid Potentials	17
	Simona Achilli, Nicola Manini, Giovanni Onida and Enrico Prati	
3	Gas and Dust Dynamics During Planet Formation in HL Tau	25
	Giulia Ballabio, Giuseppe Lodato and Giovanni Dipierro	
4	Quantum Simulation of Non-Markovian Qubit Dynamics by an All-Optical Setup	37
	Claudia Benedetti, Simone Cialdi, Matteo A. C. Rossi, Bassano Vacchini, Dario Tamascelli, Stefano Olivares and Matteo G. A. Paris	
5	Innovative Instrumentation for the Study of Atmospheric Aerosol Optical Properties	47
	Vera Bernardoni, Alice C. Forello, Federico Mariani, Bruno Paroli, Marco A. C. Potenza, Alberto Pullia, Francesco Riccobono, Tiziano Sanvito, Sara Valentini, Gianluigi Valli and Roberta Vecchi	
6	Giotto Unveiled: New Developments in Imaging and Elemental Analysis Techniques for Cultural Heritage	57
	Nicola Ludwig, Letizia Bonizzoni, Michele Caccia, Francesco Cavaliere, Marco Gargano, Daniele Viganò, Christian Salvatore, Matteo Interlenghi, Marco Martini and Anna Galli	

7	Quantitative Analysis of Gold Nano-aggregates by Combining Electron and Probe Microscopy Techniques	67
	Francesca Borghi, Matteo Mirigliano, Paolo Milani and Alessandro Podestà	
8	The Discovery of the Higgs Boson and Most Significant Measurements at the LHC	81
	Marcello Fanti	
9	Spectroscopy of Adsorbates and the Role of Interfacial Interactions	91
	Guido Fratesi, Elena Molteni and Giovanni Onida	
10	The Founder of the First Institute of Physics of the Milan University: Aldo Pontremoli, a Physicist's Life Between Adventure and Institutions	105
	Leonardo Gariboldi	
11	The Role of Monitoring Time and Detectors Efficiencies in Time-Continuous Quantum Magnetometry	127
	Francesco Albarelli, Matteo A. C. Rossi, Matteo G. A. Paris and Marco G. Genoni	
12	Recent Developments and Perspectives in Nuclear Structure by γ and Particle Spectroscopy	141
	S. Leoni, A. Bracco, F. Camera, A. Guglielmetti, G. Benzoni, B. Blasi, C. Boiano, Pier Francesco Bortignon, S. Bottoni, S. Brambilla, S. Capra, G. Colò, F. C. L. Crespi, A. Mentana, B. Million, A. Pullia, S. Riboldi, E. Vigezzi and O. Wieland	
13	Planet Formation in the ALMA Era	155
	Giuseppe Lodato, Benedetta Veronesi, Maria Giulia Ubeira Gabellini, Enrico Ragusa, Alessia Franchini, Giovanni Dipierro, Hossam Aly and Daniel J. Price	
14	Generation and Active Control of Coherent Structures in Partially-Neutralized Magnetized Plasmas	169
	Giancarlo Maero, Roberto Pozzoli, Massimiliano Romé, Beatrice Achilli, Nicola Cantini, Emilio Villa, Francesco Cavaliere and Daniele Viganò	
15	Feeding Genetic Heterogeneity via a Smart Mutation Operator in the Memetic Phase Retrieval Approach	181
	Marta Mauri, Davide Emilio Galli and Alessandro Colombo	
16	Jack on a Devil's Staircase	193
	Andrea Di Gioacchino, Marco Gherardi, Luca Guido Molinari and Pietro Rotondo	

17 A Study of the Phase Diagram of Symmetric Binary Gaussian Mixtures 209
 Davide Pini

18 Tutorial: γ_5 in Dimensional Regularization in Dimensional Regularization 221
 Mario Raciti

19 The Neutrino: Looking Through Its Experimental World 235
 Alessandra Carlotta Re

20 Towards a Universal Nuclear Structure Model 243
 X. Roca-Maza, Pier Francesco Bortignon, G. Colò, Y. F. Niu and E. Vigezzi

21 Recent Advances in Scintillating Optical Fibre Dosimeters 253
 Ivan Veronese, Marie Claire Cantone, Salvatore Gallo, Cristina De Mattia, Eduardo d’Ippolito, Nicola Ludwig, Marco Gargano, Simone Cialdi, Stefano Latorre, Norberto Chiodini, Mauro Fasoli, Federico Moretti, Eleonora Mones, Gianfranco Loi and Anna Vedda

22 R&D on Electronic Devices and Circuits for the HL-LHC 263
 Alberto Stabile, Attilio Andreatza, Mauro Citterio, Luca Frontini, Valentino Liberali, Chiara Meroni and Jafar Shojaii

23 Precise Perturbative QCD Predictions for Large Hadron Collider Physics 273
 Giancarlo Ferrera

24 LASA Cs₂Te Photocathodes: The Electron Source for XFELs 281
 Carlo Pagani, Paolo Michelato, Laura Monaco and Daniele Sertore

25 LASA Superconducting RF Cavities for Particle Accelerators 293
 Carlo Pagani, Andrea Bellandi, Michele Bertucci, Andrea Bignami, Angelo Bosotti, Jin Fang Chen, Paolo Michelato, Laura Monaco, Rocco Paparella, Daniele Sertore, Cecilia Maiano, Paolo Pierini and Saeid Pirani

Contributors

Beatrice Achilli Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Simona Achilli Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Francesco Albarelli Quantum Technology Lab, Dipartimento di Fisica “Aldo Pontremoli”, Università Degli Studi di Milano, Milan, Italy

Hossam Aly Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Attilio Andreazza Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; Istituto Nazionale di Fisica Nucleare, Milan, Italy

Giulia Ballabio Department of Physics and Astrophysics, University of Leicester, Leicester, UK; Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

J. Bel Aix–Marseille Université, Marseille, France

Andrea Bellandi INFN - LASA, Segrate (Milano), Italy

Claudia Benedetti Quantum Technology Lab, Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

G. Benzoni INFN, Sezione di Milano, Milan, Italy

Vera Bernardoni Department of Physics “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN-Milan, Milan, Italy

Michele Bertucci INFN - LASA, Segrate (Milano), Italy

D. Bianchi Institute for Cosmology and Gravitation, University of Portsmouth, Portsmouth, UK

Andrea Bignami INFN - LASA, Segrate (Milano), Italy

B. Blasi INFN, Sezione di Milano, Milan, Italy

C. Boiano INFN, Sezione di Milano, Milan, Italy

Letizia Bonizzoni Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Francesca Borghi CIMaINa and Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Pier Francesco Bortignon Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN, Sezione di Milano, Milan, Italy

Angelo Bosotti INFN - LASA, Segrate (Milano), Italy

S. Bottoni Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN, Sezione di Milano, Milan, Italy

A. Bracco Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN, Sezione di Milano, Milan, Italy

S. Brambilla INFN, Sezione di Milano, Milan, Italy

Michele Caccia CNR-IFN, Milan, Italy; CNR-IBFM, Segrate, Milan, Italy

F. Camera Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN, Sezione di Milano, Milan, Italy

Nicola Cantini Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Marie Claire Cantone Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; Istituto Nazionale di Fisica Nucleare, Sezione di Milano, Milan, Italy; Dipartimento di Scienze Biomediche, Chirurgiche ed Odontoiatriche, Università degli Studi di Milano, Milan, Italy

S. Capra Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN, Sezione di Milano, Milan, Italy

C. Carbone Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INAF - Osservatorio Astronomico di Brera, Milan, Italy

Francesco Cavaliere Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Jin Fang Chen INFN - LASA, Segrate (Milano), Italy

Norberto Chiodini Dipartimento di Scienza dei Materiali, Università di Milano-Bicocca, Milan, Italy

Simone Cialdi Quantum Technology Lab, Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN, Sezione di Milano, Milan, Italy; Istituto Nazionale di Fisica Nucleare, Sezione di Milano, Milan, Italy

Mauro Citterio Istituto Nazionale di Fisica Nucleare, Milan, Italy

G. Colò Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN, Sezione di Milano, Milan, Italy

Alessandro Colombo Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

F. C. L. Crespi Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN, Sezione di Milano, Milan, Italy

Eduardo d’Ippolito Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Cristina De Mattia Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Andrea Di Gioacchino Dipartimento di Fisica, Università degli Studi di Milano and I.N.F.N. sezione di Milano, Milan, Italy

Giovanni Dipierro Department of Physics and Astrophysics, University of Leicester, Leicester, UK; Department of Physics and Astronomy, University of Leicester, Leicester, UK

Marcello Fantì Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN, Sezione di Milano, Milan, Italy

Mauro Fasoli Dipartimento di Scienza dei Materiali, Università di Milano-Bicocca, Milan, Italy

Giancarlo Ferrera Dipartimento di Fisica “Aldo Pontremoli”, Università di Milano, Milan, Italy; INFN, Sezione di Milano, Milan, Italy

Alice C. Forello Department of Physics “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN-Milan, Milan, Italy

Alessia Franchini Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Guido Fratеси Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Luca Frontini Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; Istituto Nazionale di Fisica Nucleare, Milan, Italy

Anna Galli CNR-IFN, Milan, Italy; Dipartimento di Scienza dei Materiali, Università di Milano-Bicocca, Milan, Italy

Davide Emilio Galli Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Salvatore Gallo Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; Istituto Nazionale di Fisica Nucleare, Sezione di Milano, Milan, Italy

Marco Gargano Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Leonardo Gariboldi Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Marco G. Genoni Quantum Technology Lab, Dipartimento di Fisica “Aldo Pontremoli”, Università Degli Studi di Milano, Milan, Italy

Marco Gherardi Dipartimento di Fisica, Università degli Studi di Milano and I.N.F.N. sezione di Milano, Milan, Italy

B. R. Granett Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INAF - Osservatorio Astronomico di Brera, Milan, Italy

A. Guglielmetti Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN, Sezione di Milano, Milan, Italy

L. Guzzo Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INAF - Osservatorio Astronomico di Brera, Milan, Italy

A. J. Hawken Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INAF - Osservatorio Astronomico di Brera, Milan, Italy; Aix-Marseille Université, Marseille, France

Matteo Interlenghi CNR-IBFM, Segrate, Milan, Italy

Stefano Latorre Istituto Nazionale di Fisica Nucleare, Sezione di Milano, Milan, Italy

S. Leoni Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN, Sezione di Milano, Milan, Italy

Valentino Liberali Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; Istituto Nazionale di Fisica Nucleare, Milan, Italy

Giuseppe Lodato Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Gianfranco Loi Medical Physics Department, Azienda Ospedaliera Maggiore della Carità, Novara, Italy

Nicola Ludwig Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Giancarlo Maero Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN Sezione di Milano, Milan, Italy

Cecilia Maiano ESS Accelerator Division, Lund, Sweden

Nicola Manini Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Federico Mariani Department of Physics “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Marco Martini Dipartimento di Scienza dei Materiali, Università di Milano-Bicocca, Milan, Italy

Marta Mauri Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

A. Mentana Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN, Sezione di Milano, Milan, Italy

Chiara Meroni Istituto Nazionale di Fisica Nucleare, Milan, Italy

Paolo Michelato INFN - LASA, Segrate (Milano), Italy

Paolo Milani CIMaIna and Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

B. Million INFN, Sezione di Milano, Milan, Italy

Matteo Mirigliano CIMaIna and Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

F. G. Mohammad Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INAF - Osservatorio Astronomico di Brera, Milan, Italy

Luca Guido Molinari Dipartimento di Fisica, Università degli Studi di Milano and I.N.F.N. sezione di Milano, Milan, Italy

Elena Molteni Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Laura Monaco INFN - LASA, Segrate (Milano), Italy

Eleonora Mones Medical Physics Department, Azienda Ospedaliera Maggiore della Carità, Novara, Italy

Federico Moretti Dipartimento di Scienza dei Materiali, Università di Milano-Bicocca, Milan, Italy; Lawrence Berkeley National Laboratory, Berkeley, USA

Y. F. Niu ELI-NP, Horia Hulubei National Institute for Physics and Nuclear Engineering, Bucharest-Magurele, Romania; School of Nuclear Science and Technology, Lanzhou University, Lanzhou, China

Stefano Olivares Quantum Technology Lab, Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Giovanni Onida Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Carlo Pagani Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milano, Italy; INFN - LASA, Segrate (Milano), Italy

Rocco Paparella INFN - LASA, Segrate (Milano), Italy

Matteo G. A. Paris Quantum Technology Lab Dipartimento di Fisica “Aldo Pontremoli”, Università Degli Studi di Milano, Milan, Italy

Bruno Paroli Department of Physics “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN-Milan, Milan, Italy

A. Pezzotta Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INAF - Osservatorio Astronomico di Brera, Milan, Italy; Institute of Space Sciences (IEEC-CSIC), Barcelona, Spain

Paolo Pierini INFN - LASA, Segrate (Milano), Italy; ESS Accelerator Division, Lund, Sweden

Davide Pini Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Saeid Pirani INFN - LASA, Segrate (Milano), Italy; ESS Accelerator Division, Lund, Sweden

Alessandro Podestà CIMaINa and Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Marco A. C. Potenza Department of Physics “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN-Milan, Milan, Italy

Roberto Pozzoli Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN Sezione di Milano, Milan, Italy

Enrico Prati Istituto di Fotonica e Nanotecnologie, Consiglio Nazionale delle Ricerche, Milan, Italy

Daniel J. Price Monash Centre for Astrophysics (MoCA) and School of Physics and Astronomy, Monash University, Melbourne, Australia

A. Pullia Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN, Sezione di Milano, Milan, Italy

Alberto Pullia Department of Physics “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN-Milano, Milan, Italy

Mario Raciti Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN-Sezione di Milano, Milan, Italy

Enrico Ragusa Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Alessandra Carlotta Re Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN, Sezione di Milano, Milan, Italy

S. Riboldi Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN, Sezione di Milano, Milan, Italy

Francesco Riccobono Joint Research Centre, European Commission, Ispra, Italy

X. Roca-Maza Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN, Sezione di Milano, Milan, Italy

Massimiliano Romé Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN Sezione di Milano, Milan, Italy

Matteo A. C. Rossi Quantum Technology Lab, Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

S. Rota INAF - IASF Milano, Milan, Italy

Pietro Rotondo University of Nottingham, Nottingham, UK

Christian Salvatore CNR-IBFM, Segrate, Milan, Italy

Tiziano Sanvito Department of Physics “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; EOS s.r.l., Milan, Italy

Daniele Sertore INFN - LASA, Segrate (Milano), Italy

Jafar Shojaii Istituto Nazionale di Fisica Nucleare, Milan, Italy; The University of Melbourne, Parkville, VIC, Australia

Alberto Stabile Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; Istituto Nazionale di Fisica Nucleare, Milan, Italy

Dario Tamascelli Quantum Technology Lab, Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Maria Giulia Ubeira Gabellini Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; European Southern Observatory (ESO), Garching, Germany

Bassano Vacchini Quantum Technology Lab, Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN, Sezione di Milano, Milan, Italy

Sara Valentini Department of Physics “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN-Milan, Milan, Italy

Gianluigi Valli Department of Physics “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN-Milan, Milan, Italy

Roberta Vecchi Department of Physics “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INFN-Milan, Milan, Italy

Anna Vedda Dipartimento di Scienza dei Materiali, Università di Milano-Bicocca, Milan, Italy

Ivan Veronese Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; Istituto Nazionale di Fisica Nucleare, Sezione di Milano, Milan, Italy

Benedetta Veronesi Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

Daniele Viganò Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

E. Vigezzi INFN, Sezione di Milano, Milan, Italy

Emilio Villa Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

O. Wieland INFN, Sezione di Milano, Milan, Italy

M. Zennaro Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy; INAF - Osservatorio Astronomico di Brera, Milan, Italy

Introduction

The *Congress of the Department of Physics* has been a two-day event held in Milan on June 28–29, 2017. The Congress was mostly aimed at presenting a snapshot of the research done in our Department to the academic community, the media, and the public at large. Policy-makers and authorities were also invited. The workshop has been also instrumental in summarizing our activities and to strengthen interdisciplinary collaborations among the different areas of research within the Department, and among the members of the Department and of other communities.

We received several excellent submissions, and this fact made it rather easy to distill a program of high scientific level and, in turn, to edit the book of proceedings. Overall, serving as members of the Scientific Committee and editors of the present volume has been a fulfilling experience for all of us.

Our Department has a natural attitude toward collaborative research and interdisciplinary topics. This nature will be a key feature in the next future, while the Department of Physics, recently named after Aldo Pontremoli, will move to the new Campus at the EXPO2015 site, together with the other scientific departments.

The Congress has been attended by many researchers working in the Milan research area, and by other interested scholars as well, including students, high school teachers, and collaborators. The Congress took place in Aula Chisini at the Department of Mathematics *Federigo Enriques* of the University of Milan. We thank our colleagues for their hospitality and their logistic support.

The Congress lasted two days with a total of 25 invited presentations in eight sessions, plus two poster sessions where about 70 presentations were displayed. This volume contains contributions linked to most of the oral presentations and to a selection of the poster presentations, including the winners of the poster awards: Giulia Ballabio, Claudia Benedetti, and Andrea Merli.

We asked the contributors to provide an introduction to their fields, together with a brief account of their recent results. It has been a natural choice, following the format of the Congress, which was aimed at being a *science camp* rather than a sequence of talks, to break from the traditional conference format, which are often showcases of career-long investigations.

The overall goal of this volume is to engage the creativity of both senior and early-stage scientists and create new scientific connections, fostering critical thinking and collaborations. For these very reasons, we did not impose any constraints on contributors and you will see rather diverse styles and formats, a feature which we value as a plus. In particular, among the 25 papers that constitute this volume, you will find research papers, the most part, few tutorials, and one historical paper, by L. Gariboldi, about the life and the work of Aldo Pontremoli, who founded and directed the Institute of Advanced Physics at the University of Milan starting from 1924 until his presumed death in May 1928.

Of course, no proceedings volume may revive the whole experience of a Congress with its interactive talks and its very active poster sessions. We hope that, according to its title, the present proceedings volume will be at least successful in providing a snapshot of current research at Physics Department *Aldo Pontremoli*, and to stimulate a scientific debate as the Science Campus at the EXPO2015 site of Milan is being planned and developed. We also hope that the present material will foster additional debates on science-related topics like science communication, data visualization, open access, and social media for science.

We would like to thank all the authors for their contributions, the referees for their time and their thoroughness, and Sabine Lehr at Springer office for taking care of all the editorial and administrative matters.

Milan, Italy
March 2018

Pier Francesco Bortignon
Giuseppe Lodato
Emanuela Meroni
Matteo G. A. Paris
Alessandro Vicini

Chapter 1

Measuring the Universe with Galaxy Redshift Surveys



L. Guzzo, J. Bel, D. Bianchi, C. Carbone, B. R. Granett, A. J. Hawken,
F. G. Mohammad, A. Pezzotta, S. Rota and M. Zennaro

Abstract Galaxy redshift surveys are one of the pillars of the current standard cosmological model and remain a key tool in the experimental effort to understand the origin of cosmic acceleration. To this end, the next generation of surveys aim at achieving sub-percent precision in the measurement of the equation of state of dark energy $w(z)$ and the growth rate of structure $f(z)$. This however requires comparable control over systematic errors, stressing the need for improved modelling methods. In this paper we review a few specific highlights of the work done in this direction by the *Darklight* project (<http://darklight.fisica.unimi.it>). Supported by an ERC Advanced Grant, *Darklight* has been developing novel techniques and applying them to numerical simulations and to the new redshift survey data of the VIPERS survey. We focus in particular on: (a) advances on estimating the growth rate of structure from redshift-space distortions; (b) parameter estimation through global Bayesian reconstruction of the density field from survey data; (c) impact of massive neutrinos on large-scale structure measurements. Overall, *Darklight* is paving the way for forthcoming high-precision experiments, such as *Euclid*, the next ESA cosmological mission.

L. Guzzo (✉) · C. Carbone · B. R. Granett · A. J. Hawken
F. G. Mohammad · A. Pezzotta · M. Zennaro
Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano,
Milan, Italy
e-mail: luigi.guzzo@unimi.it

L. Guzzo · C. Carbone · B. R. Granett
A. J. Hawken · F. G. Mohammad · A. Pezzotta · M. Zennaro
INAF - Osservatorio Astronomico di Brera, Milan, Italy

J. Bel · A. J. Hawken
Aix-Marseille Université, Marseille, France

D. Bianchi
Institute for Cosmology and Gravitation, University of Portsmouth, Portsmouth, UK

A. Pezzotta
Institute of Space Sciences (IEEC-CSIC), Barcelona, Spain

S. Rota
INAF - IASF Milano, Milan, Italy

1.1 Introduction

A major achievement in cosmology over the 20th century has been the detailed reconstruction of the large-scale structure of the Universe around us. Started in the 1970s, these studies developed over the following decades into the industry of *redshift surveys*, beautifully exemplified by the Sloan Digital Sky Survey (SDSS) in its various incarnations (e.g. [1]). These maps have covered in detail our “local” Universe (i.e. redshifts $z < 0.2$) and only recently we started exploring comparable volumes at larger redshifts, where the evolution of galaxies and structure over time can be detected (see e.g. [2]). Figure 1.1 shows a montage using data from some of

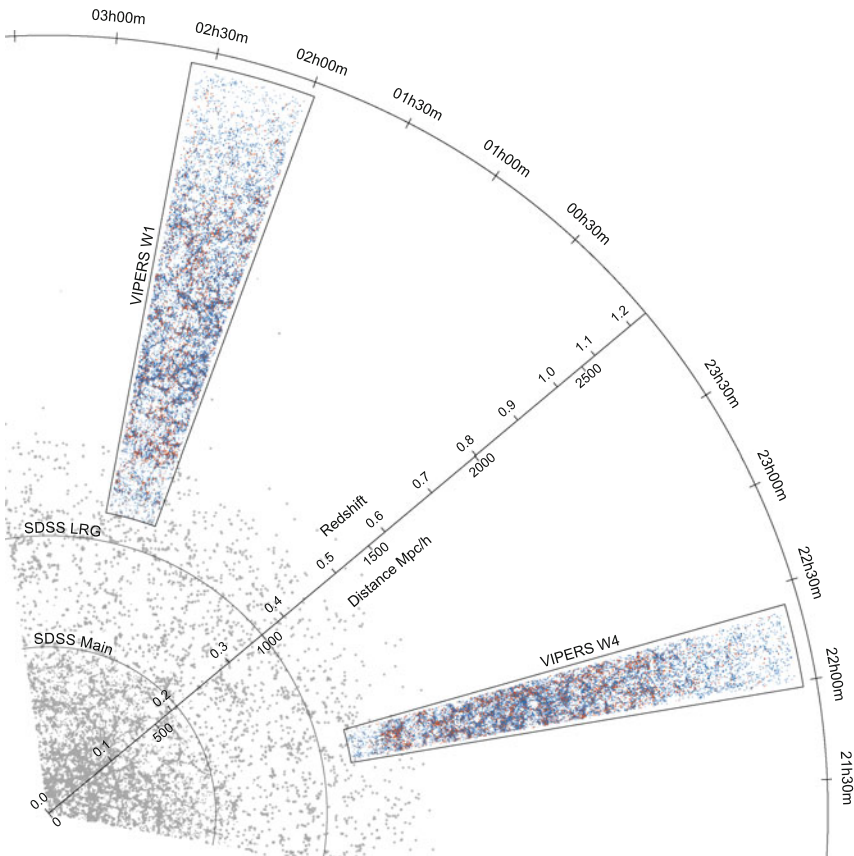


Fig. 1.1 Combined “cone diagram” of the large-scale distribution of galaxies from different surveys, out to $z = 1$. The plot includes the recently completed, deep VIPERS survey [3–5] and two sub-samples of the Sloan Digital Sky Survey (SDSS) (main sample and Luminous Red Galaxy (LRG) sample) at lower redshift [6, 7]. The plotted slices here are 4 and 2 degree-thick for the SDSS and VIPERS data, respectively

these surveys, providing a visual impression of the now well-established sponge-like topology of the large-scale galaxy distribution and how it stretches back into the younger Universe.

In addition to their purely cartographic beauty, these maps provide a quantitative test of the theories of structure formation and of the Universe composition. Statistical measurements of the observed galaxy distribution represent in fact one of the experimental pillars upon which the current “standard” model of cosmology is built. Let us define the matter over-density (or fluctuation) field, with respect to the mean density, as $\delta(\mathbf{x}) \equiv (\rho(\mathbf{x}) - \bar{\rho})/\bar{\rho}$; this can be described in terms of Fourier harmonic components as

$$\delta(\mathbf{k}) = \int_V \delta(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} d^3\mathbf{x}, \quad (1.1)$$

where V is the volume considered. The power spectrum $P(\mathbf{k})$ is then defined by the variance of the Fourier modes:

$$\langle \delta(\mathbf{k})\delta^*(\mathbf{k}') \rangle = (2\pi)^3 P(\mathbf{k})\delta_{\text{D}}(\mathbf{k} - \mathbf{k}'). \quad (1.2)$$

The observed number density of galaxies $n_g(\mathbf{x})$ is related to the matter fluctuation field through the *bias parameter* b by

$$n_g = \bar{n} (1 + b\delta), \quad (1.3)$$

which corresponds to assuming that $\delta_g = b\delta$. This linear and scale-independent relation provides an accurate description of galaxy clustering at large scales, although it breaks down in the quasi-linear regime below scales of $\sim 10 h^{-1}$ Mpc [8]. In general, b depends on galaxy properties, as we shall discuss in more detail in Sect. 1.3. From the hypothesis of linear bias, it descends that $P_{gg}(k) = b^2 P(k)$, where $P_{gg}(k)$ is the observed galaxy-galaxy power spectrum. This connection allows us to use measurements of $P_{gg}(k)$ to constrain the values of cosmological parameters that regulate the shape of $P(k)$. Figure 1.2 [9] shows an example of such measurements: the left panel plots four estimates of the power spectrum $P(k)$ (more precisely, its monopole, i.e. the average of $P(\mathbf{k})$ over spherical shells) obtained at $0.6 < z < 1.1$ from the VIPERS survey data of Fig. 1.1 (see also Sect. 1.2.2). In the central and right panels, we show the posterior distribution of the mean density of matter Ω_m and the baryon fraction f_B from a combined likelihood analysis of the four measurements; these are compared to similar estimates from other surveys and from the Planck CMB anisotropy constraints [10]. More precisely, the galaxy power spectrum shape on large scales probes the combination $\Omega_M h$, where $h = H_o/100$. Such comparisons provide us with important tests of the Λ CDM model, with the $z \sim 1$ estimate from VIPERS straddling Planck and local measurements.

If one goes beyond the simple shape of angle-averaged quantities, two-point statistics of the galaxy distribution contain further powerful information, which is key to understanding the origin of the mysterious acceleration of cosmic expansion discovered less than twenty years ago [11, 12]. First, tiny “baryonic wiggles” in

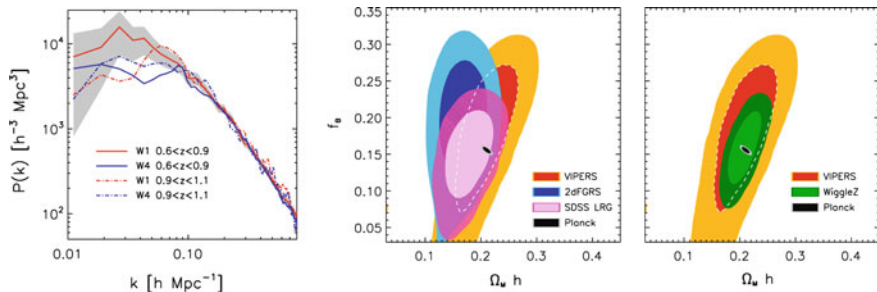


Fig. 1.2 *Left*: Four independent estimates of the power spectrum of the galaxy distribution at $0.6 < z < 1.1$ from the VIPERS galaxy survey. The four curves correspond to two redshift bins for the two separated fields W1 and W4, which have slightly different window functions (i.e. size and geometry). *Center and Right*: VIPERS constraints on the mean total matter density Ω_M times the normalised Hubble constant $h = H_o/100$ and the baryonic fraction, $f_b = \Omega_b/\Omega_M$, compared to similar measurements from surveys at low (center) and high redshift (right), plus Planck. See [9] for details

the shape of the power spectrum define a specific, well known comoving spatial scale, corresponding to the sound horizon scale at the epoch when baryons were dragged into the pre-existing dark-matter potential wells. In fact, it turns out that there are enough baryons in the cosmic mixture to influence the dominant dark-matter fluctuations [7, 13] and leave in the galaxy distribution a visible signature of the pre-recombination acoustic oscillations in the baryon-radiation plasma. Known as Baryonic Acoustic Oscillations (BAO), these features provide us with a formidable standard ruler to measure the expansion history of the Universe $H(z)$, complementary to what can be done using Type Ia supernovae as standard candles (see e.g. [14] for the latest measurements from the SDSS-BOSS sample).

Secondly, the observed redshift maps are distorted by the contribution of peculiar velocities that cannot be separated from the cosmological redshift. This introduces a measurable anisotropy in our clustering statistics, what we call Redshift Space Distortions (RSD), an effect that provides us with a powerful way to probe the *growth rate of structure* f . This key information can break the degeneracy on whether the observed expansion history is due to the presence of the extra contribution of a cosmological constant (or dark energy) in Einstein’s equations or rather require a more radical modification of gravity theory. While RSD were first described in the 1980s [15, 16]), their potential in the context of understanding the origin of cosmic acceleration was fully recognized only recently [17]; nowadays they are considered one of the potentially most powerful “dark energy tests” expected from the next generation of cosmological surveys, as in particular the ESA mission *Euclid* [18], of which the Milan group is one of the original founders.

1.2 Measuring the Growth Rate of Structure from RSD

1.2.1 Improved Models of Redshift-Space Distortions

Translating galaxy clustering observations into precise and accurate measurements of the key cosmological parameters, however, requires modelling the effects of non-linear evolution, galaxy bias (i.e. how galaxies trace mass) and redshift-space distortions themselves. The interest in RSD precision measurements stimulated work to verify the accuracy of these measurements [19, 20]. Early estimates—focused essentially on measuring Ω_M , given that in the context of General Relativity $f \simeq \Omega_M^{0.55}$ (e.g. [21])—adopted empirical non-linear corrections to the original linear theory by Kaiser; this is the case of the so-called “dispersion model” [22], which in terms of the power spectrum of density fluctuations is expressed as

$$P^s(k, \mu) = D(k\mu\sigma_{12}) \left(1 + \beta\mu^2\right)^2 b^2 P_{\delta\delta}(k), \quad (1.4)$$

where $P^s(k, \mu)$ is the redshift-space power spectrum, which depends both on the amplitude k and the orientation $\mu = \cos(\Phi)$ of the Fourier mode with respect to the line-of-sight, $P_{\delta\delta}(k)$ is the real-space (isotropic) power spectrum of the matter fluctuation field δ and $\beta = f/b$, with f being the growth of structure and b the *linear bias* of the specific population of halos (or galaxies) used. The latter is defined as the ratio of the *rms* clustering amplitude of galaxies to that of the matter, conventionally measured in spheres of $8 \text{ h}^{-1} \text{ Mpc}$ radius, $b = \sigma_8^{\text{gal}}/\sigma_8$. For what will follow later, it is useful to note that

$$\beta = \frac{f}{b} = f \frac{\sigma_8}{\sigma_8^{\text{gal}}}, \quad (1.5)$$

can be recast as

$$\beta\sigma_8^{\text{gal}} = f\sigma_8, \quad (1.6)$$

which combines two directly measurable quantities to the left, showing that what we actually measure is the combination of the growth rate and the *rms* amplitude of clustering, $f\sigma_8$. This is what nowadays is customarily plotted when presenting measurements of the growth rate from redshift surveys (e.g. Fig. 1.8).

Going back to (1.4), the term $D(k\mu\sigma_{12})$ is usually either a Lorentzian or a Gaussian function, empirically introducing a nonlinear damping to the Kaiser linear amplification, with the Lorentzian (corresponding to an exponential in configuration space) normally providing a better fit to the galaxy data [23]. This term is regulated by a second free parameter, σ_{12} , which corresponds to an effective (scale-independent) line-of-sight pairwise velocity dispersion. Figure 1.3 (from [20]), shows how estimates of β using the dispersion model can be plagued by systematic errors as large as 10%, depending on the kind of galaxies (here dark matter halos) used. With the next

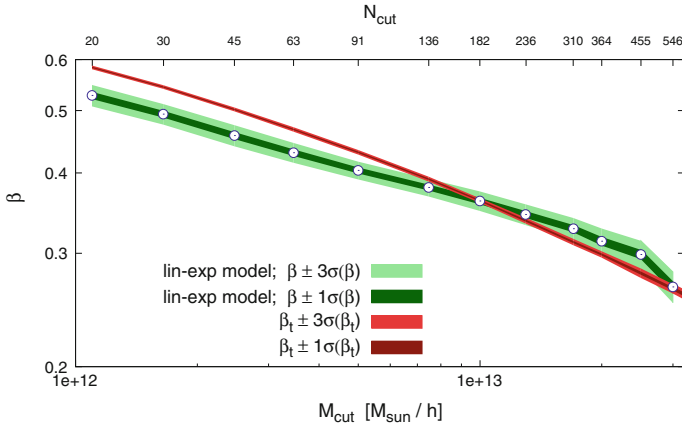


Fig. 1.3 Systematic differences in the measured values of the RSD distortion parameter β (dots with green 1- and 2- σ error bands) with respect to the expected value (thinner red band, including theoretical uncertainties). Measurements are performed for catalogues of dark-matter halos with increasing threshold mass, built from an n-body simulation [20]

generation of surveys aiming at 1% precision by collecting several tens of millions of redshifts, such a level of systematic errors is clearly unacceptable.

Exploring how to achieve this overall goal by optimising measurements of galaxy clustering and RSD, has been one of the main goals of the *Darklight* project, supported by an ERC Advanced Grant awarded in 2012. *Darklight* focused on developing new techniques, testing them on simulated samples, and then applying them to the new data from the VIMOS Public Extragalactic Redshift Survey (VIPERS), which was built in parallel.

After assessing the limitations of existing RSD models [20, 24] the first goal of *Darklight* has been to develop refined theoretical descriptions. This work followed two branches: one, starting from first principles, was based on revisiting the so-called *streaming model* approach; the second, more pragmatic, aimed at refining the application to real data of the best models available at the time, as in particular the “TNS” model [25]. Such more “data oriented” line of development also included exploring the advantages of specific tracers of large-structure in reducing the impact of non-linear effects.

The first approach [26] focused on the so-called *streaming model* [27], which in the more general formulation by Scoccimarro [28] (see also [29]), describes the two-point correlation function in redshift space $\xi_S(s_\perp, s_\parallel)$ as a function of its real-space counterpart $\xi_R(r)$

$$1 + \xi_S(s_\perp, s_\parallel) = \int dr_\parallel [1 + \xi_R(r)] \mathcal{P}(r_\parallel - s_\parallel | \mathbf{r}). \quad (1.7)$$

Here quantities noted with \perp and \parallel correspond to the components of the pair separation—in redshift or real space—respectively perpendicular and parallel to the

line of sight, with $r^2 = r_{\parallel}^2 + r_{\perp}^2$ and $r_{\perp} = s_{\perp}$. The interest in the streaming model is that this expression is exact: knowing the form of the pairwise velocity distribution function $\mathcal{P}(v_{\parallel}|\mathbf{r}) = \mathcal{P}(r_{\parallel} - s_{\parallel}|\mathbf{r})$ at any separation \mathbf{r} , a full mapping of real- to redshift-space correlations is provided. The problem is that this is a virtually infinite family of distribution functions.

The essential question addressed in [26] has been whether a sufficiently accurate description of this family (and thus of RSD) is still possible with a reduced number of degrees of freedom. It is found that, at a given galaxy separation \mathbf{r} , they can be described as a superposition of virtually infinite Gaussian functions, whose mean μ and dispersion σ are in turn distributed according to a bivariate Gaussian, with its own mean and covariance matrix. A recent extension of this work [30] shows that such ‘‘Gaussian–Gaussian’’ model cannot fully match the level of skewness observed at small separations, in particular when applied to catalogues of dark matter halos. They thus generalize the model by allowing for the presence of a small amount of local skewness, meaning that the velocity distribution is obtained as a superposition of quasi-Gaussian functions. In its simplest formulation, this improved model takes as input the real space correlation function and the first three velocity moments (plus two well defined nuisance parameters) and returns an accurate description of the anisotropic redshift-space two-point correlation function down to very small scales ($\sim 5 \text{ h}^{-1} \text{ Mpc}$ for dark matter particles and virtually zero for halos). To be applied to real data to estimate the growth rate of structure f , the model still needs a better theoretical and/or numerical understanding of how the velocity moments depend on f on small scale, as well as tests on mock catalogues including realistic galaxies.

The second, parallel approach followed in *Darklight* was to work on the ‘‘best’’ models existing in the literature, optimising their application to real data. The natural extensions to the dispersion model (1.4) start from the Scoccimarro [28] expression

$$P^s(k, \mu) = D(k\mu\sigma_{12}) \left(b^2 P_{\delta\delta}(k) + 2fb\mu^2 P_{\delta\theta}(k) + f^2\mu^4 P_{\theta\theta}(k) \right), \quad (1.8)$$

where $P_{\delta\theta}$ and $P_{\theta\theta}$ are respectively the so-called density-velocity divergence cross-spectrum and the velocity divergence auto-spectrum, while $P_{\delta\delta}$ is the usual matter power spectrum. If one then also accounts for the non-linear mode coupling between the density and velocity-divergence fields, two more terms arise inside the parenthesis, named $C_A(k, \mu, f, b)$ and $C_B(k, \mu, f, b)$, leading to the TNS model by Taruya and collaborators [25].

A practical problem in the application of either of these two models is that the values of $P_{\delta\theta}$ and $P_{\theta\theta}$ cannot be measured from the data. As such, they require empirical fitting functions to be calibrated using numerical simulations [31]. As part of the *Darklight* work, we used the DEMNUni simulations (see Sect. 1.4) to derive improved fitting functions in different cosmologies [32]:

$$P_{\delta\theta}(k) = \left(P_{\delta\delta}(k) P^{\text{lin}}(k) e^{-k/k^*} \right)^{\frac{1}{2}}, \quad (1.9)$$

$$P_{\theta\theta}(k) = P^{\text{lin}}(k)e^{-k/k^*}, \quad (1.10)$$

where $P^{\text{lin}}(k)$ is the linear matter power spectrum and k^* is a parameter representing the typical damping scale of the velocity power spectra, which is well described as $1/k^* = p_1\sigma_8^{p_2}$, where p_1 , p_2 are the only two parameters that need to be calibrated from the simulations. These forms for $P_{\delta\delta}$ and $P_{\theta\theta}$ have valuable, physically motivated properties: they naturally converge to $P_{\delta\delta}(k)$ in the linear regime, including a dependence on redshift through $\sigma_8(z)$. They represent a significant improvement over previous implementations of the Scoccimarro and TNS models and allowed us to extend their application to smaller scales and to the high redshifts covered by VIPERS.

1.2.2 Application to Real Data: Optimising the Samples

The performance, in terms of systematic error, of any RSD model when applied to real data does not depend only on the quality of the model itself. The kind of tracers of the density and velocity field that are used, significantly enhance or reduce some of the effects we are trying to model and correct. This means that, in principle, we may be able to identify specific sub-samples of galaxies for which the needed non-linear corrections to RSD models are intrinsically smaller. This could be an alternative to making our models more and more complex, as it happens for the full galaxy population.

Such an approach becomes feasible if the available galaxy survey was constructed with a broad selection function and supplemented by extensive ancillary information (e.g. multi-band photometry, from which spectral energy distributions, colours, stellar masses, etc. can be obtained). This allows a wide space in galaxy physical properties to be explored, experimenting with clustering and RSD measurements using different classes of tracers (and their combination), as e.g. red versus blue galaxies, groups, clusters. This is the case, for example, of the Sloan Digital Sky Survey main sample [6]. The VIMOS Public Extragalactic Redshift Survey (VIPERS) [3] was designed with the idea of extending this concept to $z \sim 1$, i.e. when the Universe was around half its current age, providing *Darklight* with a state-of-the-art playground.

VIPERS is a new statistically complete redshift survey, constructed between 2008 and 2016 as one of the ‘‘ESO Large Programmes’’, exploiting the unique capabilities of the VIMOS multi-object spectrograph at the Very Large Telescope (VLT) [5]. It has secured redshifts for 86,775 galaxies with magnitude $i_{AB} \leq 22.5$ (out of 97,714 spectra) over a total area of 23.6 square degrees, tiled with a mosaic of 288 VIMOS pointings. Target galaxies were selected from the two fields (W1 and W4) of the Canada-France-Hawaii Telescope Legacy Survey Wide catalogue (CFHTLS-Wide), benefiting of its excellent image quality and photometry in five bands (*ugriz*).¹ The

¹<http://www.cadc-ccda.hia-ihp.nrc-cnrc.gc.ca/en/cfht>.

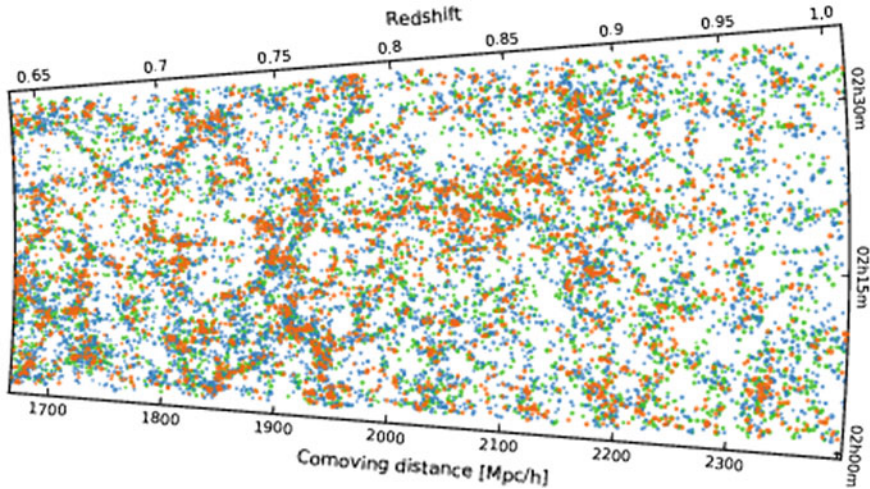


Fig. 1.4 A zoom into the central part of the W1 VIPERS region. Galaxies are described by dots, whose size is proportional to the B -band luminosity of the galaxy and whose colour corresponds to its actual (UB) restframe colour. Note the clear colour-density relation, for the first time seen so clearly at these redshifts, with red early-type galaxies tracing the backbone of structure and blue/green star-forming objects filling the more peripheral lower-density regions

survey concentrates over the range $0.5 < z < 1.2$, thanks to a robust colour pre-selection that excluded lower- z targets, nearly doubling in this way the sampling density achieved by VIMOS within the redshift of interest [3]. This set-up produces a combination of dense sampling ($>40\%$) and large volume ($\sim 5 \times 10^7 \text{ h}^{-3} \text{ Mpc}^3$), which is unique for these redshifts and allows studies of large-scale structure and galaxy evolution to be performed on equal statistical footing with state-of-the-art surveys of the local $z < 0.2$ Universe (see Fig. 1.1). Sparser samples like the SDSS LRG, BOSS [14] or Wigglez [33] surveys allow for much larger volumes to be probed and are excellent to measure large-scale features as Baryonic Acoustic Oscillations. However, they include a very specific, limited sample of the overall galaxy population and (by design) fail to register the details of the underlying nonlinear structure. The rich content of information of VIPERS can be further appreciated in Fig. 1.4, where the connection between galaxy colours and large-scale structure is readily visible by eye. VIPERS released publicly its final catalogue and a series of new scientific results in November 2016. More details on the survey construction and the properties of the sample can be found in [3–5].

Figure 1.5 shows two measurements of the anisotropic two-point correlation function in redshift space (i.e. what is called $\xi_S(s_\perp, s_\parallel)$ in (1.7)); here $r_p = s_\perp$ and $\pi = s_\parallel$), using the VIPERS data. In this case the sample has been split into two classes, i.e. blue and red galaxies, defined on the basis of their rest-frame ($U - V$) photometric colour (see [34] for details). The signature of the linear streaming motions produced by the growth of structure is evident in the overall flattening of the contours along the line-of-sight direction (π). These plots also show how blue galaxies (left) are less

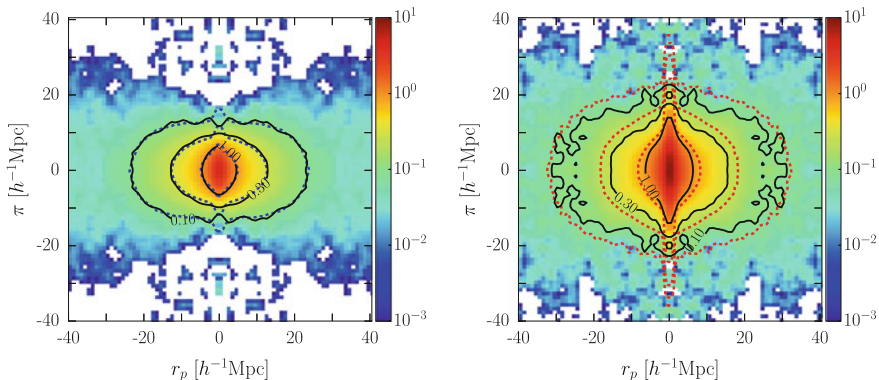


Fig. 1.5 Estimate of the redshift-space two-point correlation functions from the VIPERS survey, splitting the sample into blue (left) and red (right) galaxies (colour scale and solid contours), compared to measurements from a set of mock samples (dashed lines). Blue galaxies show reduced stretching along the line-of-sight (π) direction, indicating lower contribution by non-streaming velocities, which are the most difficult to account for in the extraction of the linear component and the growth rate of structure f [34]

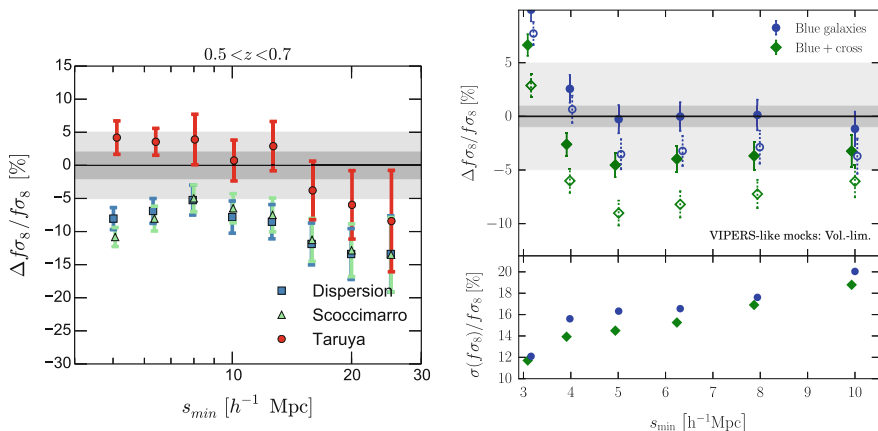


Fig. 1.6 Systematic errors on the growth rate parameter $f\sigma_8$ using 153 VIPERS-like mock catalogues. In both panels the abscissa correspond to the minimum scale included in the fit: the smaller s_{\min} , the more nonlinear effects are included. *Left*: improving non-linear corrections in the RSD model [23]. *Right*: improving the galaxy tracers: luminous blue galaxies yield negligible systematic errors down to $5 h^{-1}$ Mpc, even limiting non-linear corrections to the Scoccimarro extension (filled circles) of the dispersion model (open circles)

affected by small-scale nonlinear motions, i.e. those of high-velocity pairs within virialised structures. These produce the small-scale stretching of the contours along π (vertical direction), which is instead evident in the central part of the red galaxy plot on the right. For this reason, blue galaxies turn out to be better tracers of RSD, for which it is sufficient to use a simpler modelling, as shown in Fig. 1.6. When using the full galaxy population, the best performing model is the TNS by Taruya et al.

[25] (left panel), while when we limit the sample to luminous blue galaxies only, it is sufficient to use the simpler nonlinear corrections by Scoccimarro [28] (filled circles, right panel); open circles correspond to the simplest model, i.e. the standard dispersion model [22], which is not sufficient even in this case. See [34] for details.

1.2.3 RSD from Galaxy Outflows in Cosmic Voids

Cosmic voids, i.e. the large under-dense regions visible also in Fig. 1.1, represent an interesting new way to look at the data from galaxy redshift surveys. As loose as they may appear, over the past few years they have proved to be able to yield quantitative cosmological constraints on the growth of structure. Indeed, growth-induced galaxy peculiar velocities tend to outflow radially from voids, which leaves a specific mark in the observed void-galaxy cross-correlation function (see e.g. [35]). The dense sampling of VIPERS makes it excellent for looking for cosmic voids at high redshift. Figure 1.7 shows an example of how a catalogue of voids was constructed from these data [36].

The *Darklight* contribution to this new research path has been presented recently [37]. By modelling the void-galaxy cross-correlation function of VIPERS, a further complementary measurement of the growth rate of structure has been obtained [37]. This value is plotted in Fig. 1.8, which provides a summary of all VIPERS estimates, plotted in the customary form $f\sigma_8$ (see Sect. 1.2.1 for details). The figure also includes one further measurement, based on a joint analysis of RSD and galaxy-galaxy lensing [38], which has not been discussed here. In addition, one more analysis is in progress, based on the linearisation technique called “clipping” [39].

Such a multifaceted approach to estimating the growth rate of structure clearly represents an important cross-check of residual systematic errors in each single technique. We stress again how this has been made possible thanks to the broad “information content” of the VIPERS survey, which provides us with an optimal compromise (for these redshifts) between a large volume, a high sampling rate and extensive information on galaxy physical properties.

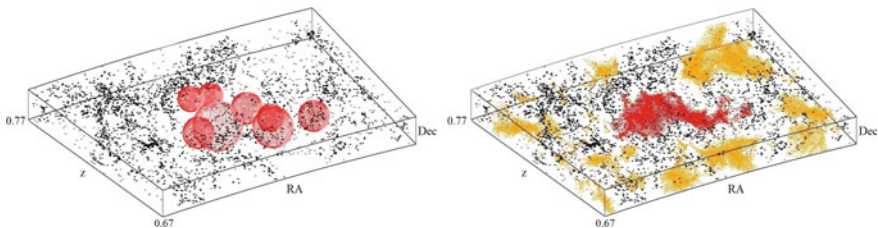


Fig. 1.7 Example of definition and search for “voids”, as performed in VIPERS. *Left*: the spherical void regions that make up the largest void in one of the VIPERS fields. *Right*: in red, the centres of all overlapping significant spheres defining the same low-density region; other void regions within this volume are shown in orange [36]

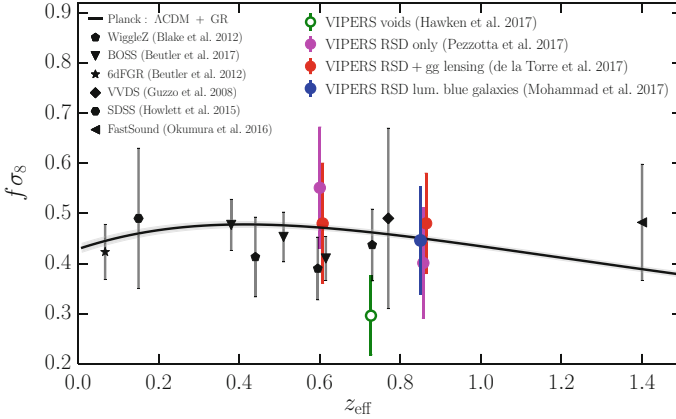


Fig. 1.8 Collection of all VIPERS estimates of the growth rate of structure using complementary methods, compared to literature results [17, 40–44], and the prediction of the standard cosmological model with current Planck parameter values and uncertainties (solid black/grey band) [10]

1.3 Optimal Methods to Derive Cosmological Parameters

The cosmological information we are interested in is encoded in the two-point statistics of the matter density field, i.e. its correlation function or, in Fourier space, its power spectrum $P_{\delta\delta}$. As we have seen in the Introduction, this is connected to the observed galaxy fluctuations as $P_{gg}(k) = b^2 P_{\delta\delta}(k)$, with $n_g = \bar{n} (1 + b\delta)$. The galaxy bias b depends in general on the galaxy properties, such as their luminosity and morphology, as well as the environment in which they are found (in groups or in isolation). Thus, in this context the bias terms are nuisance parameters that are marginalized in the analysis. However, the precision with which the measurement can be made depends very much on these parameters as they set the amplitude of the power spectrum and the effective signal-to-noise ratio.

Going beyond the standard approach to estimate cosmological parameters, as e.g. used in the $P(k)$ analysis of Fig. 1.2, in *Darklight* we have investigated and applied optimal methods given the observed constraints (luminosity function and bias). We can formulate this as a forward modelling problem through Bayes’ theorem, which tells us how the measurements relate to the model:

$$p(P_{\delta\delta}, \delta, b, \bar{n} | n_g) \propto p(n_g | P_{\delta\delta}, \delta, b, \bar{n}) p(P_{\delta\delta}, \delta, b, \bar{n}). \quad (1.11)$$

On the left-hand side, the posterior describes the joint distributions of the model parameters, here explicitly written as the density field δ , its power spectrum $P_{\delta\delta}$, the galaxy bias b and the mean number density \bar{n} , but we can generalize to the underlying cosmological parameters. The posterior is factored into the likelihood and prior terms on the right-hand side. To evaluate the posterior we must assume forms for these functions. We begin by assuming multi-variate Gaussian distributions

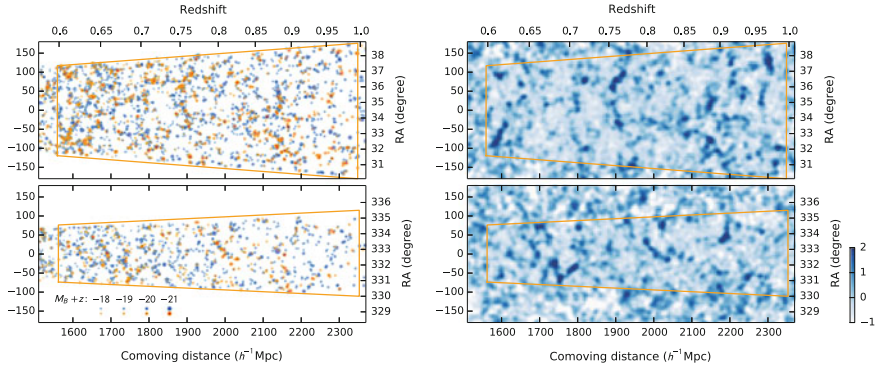


Fig. 1.9 *Left:* Each point in these slices through the VIPERS W1 (top) and W4 (bottom) fields represents a galaxy with a particular luminosity and color. *Right:* The Wiener reconstruction of the density field computed from these observations. Beyond the survey limits the density field is generated with a constrained random Gaussian field. See [45] for details

for the likelihood and priors since these forms fully encode the information contained in the power spectrum or correlation function statistics. In this limit the maximum-likelihood solution is given by the Wiener filter. In [45] we demonstrate that in this limit the solution is optimal in the sense that it minimizes the variance on the density field and power spectrum.

Figure 1.9 shows one possible reconstruction of the VIPERS density field. It represents a single step in the Monte Carlo chain used to sample the full posterior distribution as presented in [45]. In this work we characterized the full joint posterior likelihood of the density field, the matter power spectrum, RSD parameters, linear bias and luminosity function. These terms, particularly since they are estimated from a single set of observations, are correlated and the analysis naturally reveals these correlations.

A notable aspect of this analysis is that we optimally use diverse information including the luminosity function, density field and power spectrum to infer cosmological parameters and it becomes even more interesting with additional observables. We can envision simultaneous inference using cluster counts or cosmic shear. Generalizing requires putting a full dynamical model for large-scale structure in the likelihood term effectively moving the likelihood analysis to the initial conditions. Observational systematics may be naturally included as well.

1.4 A New Kid in Town: Massive Neutrinos

The non-vanishing neutrino mass, implied by the discovery of neutrino flavour oscillations, has important consequences for our analysis of the large-scale structure in the Universe. Even if sub-dominant, the neutrino contribution suppresses to some extent

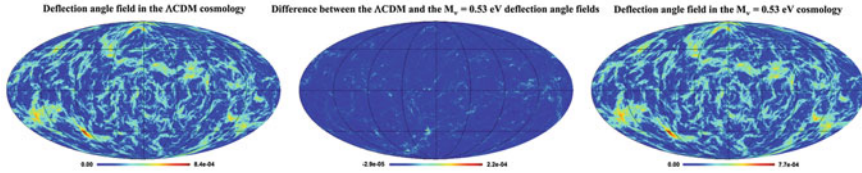


Fig. 1.10 Synthetic weak-lensing maps obtained via ray-tracing across the matter distribution in the DEMNUni simulations, for lensed sources spherically placed at $z = 1$ around the observer. The left and right maps show the amplitude of the deflection angle for pure Λ CDM and for the case with $M_\nu = 0.53$ eV total neutrino mass, respectively. The difference of the two maps is shown in the middle, and represents a change of about 6% both in the *mean* and *rms* values of the deflection angle fields between the two scenarios

the growth of fluctuations on specific scales, producing a deformation of the shape of the total matter power spectrum. Given current upper limits on the sum of the masses M_ν ($M_\nu \leq 0.16$ eV at 95% confidence [14]), the expected effect corresponds to a few percent change in the amplitude of total matter clustering. In the era of precision cosmology, neutrinos are an ingredient that cannot be neglected anymore. Conversely, future surveys like Euclid may eventually be able to obtain an estimate of the total mass of neutrinos with a precision that surpasses ground-based experiments [46]. To achieve this goal, we shall be able to: (a) describe how these effects are mapped from the matter to the galaxy power spectrum, i.e. what we measure; (b) distinguish these spectral deviations from those due to non-linear clustering, and to the presence of other possible contributions, e.g. forms of dark energy beyond the cosmological constant, like quintessence or in general an evolving equation of state of dark energy $w(z)$.

This has been addressed in *Darklight* through the “Dark Energy and Massive Neutrino Universe” (DEMNUni) simulations, a suite of fourteen large-sized N-body runs including massive neutrinos (besides cold dark matter), which have been recently completed [47]. They explore the impact on the evolution of structure of a neutrino component with three different total masses ($M_\nu = 0.17, 0.30, 0.53$ eV), including scenarios with evolving $w(z)$, according to the phenomenological form $w(z) = w_0 + w_a z / (1 + z)$.

Running these simulations required developing new techniques to account for the evolving hot dark matter component represented by neutrinos [48]. Early analyses of the whole suite show that the effects of massive neutrinos and evolving dark energy are highly degenerate (less than 2% difference) with a pure Λ CDM model, when one considers the clustering of galaxies or weak lensing observations. Disentangling these different effects will therefore represent a challenge for future galaxy surveys as Euclid and needs to be carefully addressed.

Figure 1.10 gives an example of physical effects that can be explored using these numerical experiments, showing weak-lensing maps (in terms of the amplitude of the resulting deflection angle) built via ray-tracing through the matter particle distribution of the simulations, for sources placed at redshift $z = 1$. The middle panel shows the

difference between a pure Λ CDM scenario and a model with $M_\nu = 0.53$ eV. More quantitatively, in terms of angular power spectra of the deflection field, massive neutrinos produce a scale-dependent suppression with respect to the Λ CDM case, which, on small scales, asymptotically tends towards a constant value of about 10%, 19%, 31% for $M_\nu = 0.17, 0.30, 0.53$ eV, respectively.

Acknowledgements We thank our collaborators in the VIPERS team for their contribution to building and analysing such a unique galaxy sample.

References

1. D.J. Eisenstein et al., *AJ* **142**, 72 (2011)
2. L. Guzzo, Vipers team. *The Messenger* **168**, 40 (2017)
3. L. Guzzo et al., *AAP* **566**, A108 (2014)
4. B. Garilli et al., *AAP* **562**, A23 (2014)
5. M. Scodeggio et al., *AAP* (2017), [arXiv:1611.07048](https://arxiv.org/abs/1611.07048)
6. D.G. York et al., *AJ* **120**, 1579 (2000)
7. D.J. Eisenstein et al., *ApJ* **633**, 560 (2005)
8. C. Di Porto et al., *AAP* **594**, A62 (2016)
9. S. Rota et al., *AAP* **601**, A144 (2017)
10. Planck Collaboration, et al. (2015), [arXiv:1502.01589](https://arxiv.org/abs/1502.01589)
11. A.G. Riess et al., *AJ* **116**, 1009 (1998)
12. S. Perlmutter et al., *ApJ* **517**, 565 (1999)
13. S. Cole et al., *MNRAS* **362**, 505 (2005)
14. S. Alam et al., *MNRAS* **470**, 2617 (2017)
15. M. Davis, P.J.E. Peebles, *ApJ* **267**, 465 (1983)
16. N. Kaiser, *MNRAS* **227**, 1 (1987)
17. L. Guzzo et al., *Nature* **451**, 541 (2008)
18. R. Laureijs et al. (2011), [arXiv:1110.3193](https://arxiv.org/abs/1110.3193)
19. T. Okumura, Y.P. Jing, *ApJ* **726**, 5 (2011)
20. D. Bianchi et al., *MNRAS* **427**, 2420 (2012)
21. J.A. Peacock et al., *Nature* **410**, 169 (2001)
22. J.A. Peacock, S.J. Dodds, *MNRAS* **267**, 1020 (1994)
23. A. Pezzotta et al., *AAP* **604**, A33 (2017)
24. S. de la Torre, L. Guzzo, *MNRAS* **427**, 327 (2012)
25. A. Taruya, T. Nishimichi, S. Saito, *Phys. Rev. D* **82**(6), 063522 (2010)
26. D. Bianchi, M. Chiesa, L. Guzzo, *MNRAS* **446**, 75 (2015)
27. K.B. Fisher, *ApJ* **448**, 494 (1995)
28. R. Scoccimarro, *Phys. Rev. D* **70**(8), 083007 (2004)
29. B.A. Reid et al., *MNRAS* **426**, 2719 (2012)
30. D. Bianchi, W.J. Percival, J. Bel, *MNRAS* **463**, 3783 (2016)
31. E. Jennings, C.M. Baugh, S. Pascoli, *MNRAS* **410**, 2081 (2011)
32. J. Bel et al., in preparation (2017)
33. C. Blake et al., *MNRAS* **406**, 803 (2010)
34. F.G. Mohammad et al., *ArXiv e-prints* (2017)
35. N. Hamaus, *Phys. Rev. Lett.* **117**(9), 091302 (2016)
36. D. Micheletti et al., *AAP* **570**, A106 (2014)
37. A.J. Hawken et al., *AAP* (2017), [arXiv:1611.07046](https://arxiv.org/abs/1611.07046)
38. S. de la Torre et al., *AAP* (2017), [arXiv:1612.05647](https://arxiv.org/abs/1612.05647)
39. M. Wilson et al., in preparation (2017)

40. C. Blake et al., MNRAS **425**, 405 (2012)
41. F. Beutler et al., MNRAS **466**, 2242 (2017)
42. F. Beutler et al., MNRAS **423**, 3430 (2012)
43. C. Howlett et al., MNRAS **449**, 848 (2015)
44. T. Okumura, Pub. Astr. Soc. Jpn. **68**, 38 (2016)
45. B.R. Granett et al., AAP **583**, A61 (2015)
46. C. Carbone et al., JCAP **3**, 030 (2011)
47. C. Carbone, M. Petkova, K. Dolag, JCAP **7**, 034 (2016)
48. M. Zennaro et al., MNRAS **466**, 3244 (2017)

Chapter 2

Theoretical Spectroscopy

Characterization of Deep Electronic States of Defects in Silicon via Density Functional Theory with Hybrid Potentials



Simona Achilli, Nicola Manini, Giovanni Onida and Enrico Prati

Abstract Starting with an introductory overview of the relevant aspects that characterize single-electron transport via individual dopant atom in silicon, we briefly discuss the role of theoretical calculations at different levels of approximation, as supporting tool for the interpretation of the experiments. In particular we suggest to adopt Density Functional Theory with screened exchange hybrid potential as suitable formalism to achieve an accurate description of the local geometry and excited states electronic structure of dopant-related defects. We present the results obtained for the test case of a single vacancy in silicon considering the band structure as benchmark to appreciate the advantages of the method. A complete characterization of the dopant-related levels in the gap would require to evaluate the charge transition state levels. This calculation paves the way towards a systematic study of more complex defects within the same theoretical scheme, including systems of applicative interest.

2.1 Introduction

Single atom semiconductor devices have been recently proposed as pathway toward the realization of quantum transport in nano-transistors [1, 2]. They operate exploiting the single-electron tunneling through localized states that lie within the energy gap of the semiconductor matrix. These atomic levels can be achieved through the inclusion of donors and acceptor behaving as quantum dots. Tunneling transport via individual dopant atoms in silicon was discovered accidentally and it can

S. Achilli (✉) · N. Manini · G. Onida
Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano,
Via Celoria 16, Milan, Italy
e-mail: simona.achilli@unimi.it

E. Prati
Istituto di Fotonica e Nanotecnologie, Consiglio Nazionale delle Ricerche,
Piazza Leonardo da Vinci 32, Milan, Italy

represent a detrimental process if the dopants are located randomly in the channel, making the device not reproducible [3]. Recently, high accuracy in the control of dopant position has been achieved by means of single-ion implantation technique. Different samples with conventional dopants implanted in the silicon channel exhibited quantum dot behaviour [4]. Electrical measurements of the transport properties of such devices showed indeed that controlling the gate voltage and the source-drain potential it is possible to address, at cryogenic temperature, I–V curves characterized by isolated current peaks. These features can be ascribed to single-electron tunneling via dopant states that lie below the Fermi energy of the leads for certain values of the gate voltage. Increasing the dopant concentration the Coulomb interaction between the dopants and the disorder related effects give rise to Anderson-Mott transition controlled by the temperature [5]. In highly doped samples fingerprint of the formation of Hubbard bands has been observed in the I–V curves and different thermally activated transport mechanisms have been suggested. The activation energy of hopping processes can be deduced by the Arrhenius-like behaviour of the conductance as a function of the temperature and it is of the order of MeV.

Silicon devices exploiting conventional dopants atoms (P, As) are characterized by low operating temperatures, being the electronic levels in the gap only 0.02/0.05 eV below the conduction band, leading to almost complete ionization at room temperature. Implantation of different kind of ions could reasonably give rise to deeper electronic levels and guarantee single-electron transport also at room temperature. For example, a possible route toward high-temperature operating conditions could be the implantation of Ge ion that should create, upon annealing, local defect-complexes characterized by deeper electronic states [6].

Notably, the low concentrations of implanted ions (ranging from 10^{15} to 10^{19} cm^{-3}) prevents the characterization of the samples by means of standard techniques, such as electron paramagnetic resonance (EPR) measurements [7], determining an uncertainty relatively to the local arrangement of the defect. Furthermore, the energy position of the single atomic levels in the gap is not easily obtained from the electrical measurements due to the issue in the conversion between applied gate potential and energy level position.

A complete understanding of the experimental data can take advantage from theoretical tools able to shed light on the most likely defect configuration and the relative electronic properties. In particular theoretical spectroscopy calculations allow one to determine the electronic states involved in the Hubbard band formation and the excitation energies associated to electronic transitions involving defect levels. A suitable theoretical method is expected to correctly describe the local geometry of the defect, the impurity levels in the energy gap, the charge transition levels and the wavefunction localization in the host material. Different approaches can be adopted, depending on the degree of accuracy and computational cost required. Density Functional Theory (DFT) with standard GGA-LDA approximation [8, 9] can be successfully applied to reproduce spectroscopic data of occupied states of metals and low dimensional systems [10, 11], to describe extended states [12–14], and to study the atom-surface interaction mechanisms [15–17]. Differently, the characterization of excited states properties and spatially localized electronic states suffers the intrinsic limitation of

such approximations. In the next section we will show that DFT approach with hybrid functional seems to be a viable strategy to obtain good accuracy in the results relative to excited states properties and localized electronic states, without a huge computational effort.

2.2 Theoretical Approach

As mentioned above, the study of the electronic properties of defects originated by single-atom implantation in silicon requires an adequate description of defect levels in the band gap including their wavefunction and the distribution of the related excess charge near the defect. Standard DFT calculations with local or semilocal functionals such as local density approximation (LDA) [8] or generalized gradient approximation (GGA) [9] are prone to self-interaction errors [18] leading to the wavefunction delocalization and to a bad description of silicon energy gap and of the excited states. In fact, the electronic gap obtained in the Khon-Sham scheme differs from the true quasiparticle gap (measured e.g. in photoemission and inverse photoemission experiments) due to the derivative discontinuity of the exchange-correlation (XC) energy functional [19]. A rigorous description of electronic excitations requires working in the framework of many-body perturbation theory [20]. In order to solve this issue more refined and sophisticated approaches, such as the GW method [21], should be used.

GW can be nowadays applied to quite complex systems [22–24]. However, the large supercells needed to reduce the interaction between periodic replica of the defect make the full GW approach not feasible for the system considered here. Moreover, total energies, although in principle accessible through the Galitsky-Migdal formula, are still out of the reach of the GW approach for realistic systems.

A suitable alternative is represented by the use of hybrid functionals that allow systematic and quantitative calculations of the interplay between structural and electronic properties. Hybrid functionals are defined as a one-parameter dependent admixture of exact (E_X) and GGA exchange and correlation energy:

$$E_{xc} = \alpha E_x + (1 - \alpha) E_X^{GGA} + E_C^{GGA} \quad (2.1)$$

where α is the added fraction of exact exchange which permits to correct the self-interaction error and to describe the long range Coulomb interaction.

Standard recipes for hybrid functionals prescribe a fixed portion α of exact exchange, independently of the system being investigated. This is the case of the full-range Perdew-Burke-Ernzerhof (PBE0) [25] in which α is set to 0.25 and the three parameters B3LYP functional [26]. Further improvements, especially in the description of metallic systems, has been obtained through screened-exchange hybrid functionals, such as Heyd-Scuseria-Ernzerhof (HSE06) [27], in which the exchange component is separated into a long-range and a short-range contribution.

It has been demonstrated recently that the appropriate amount of exchange for a given system can be related to one of its physical properties: the static electronic screening. Indeed in the framework of many-body perturbation theory and following the GW approach, the α parameter results equal to $1/\epsilon_\infty$ provided that microscopic components of the dielectric function are averaged out over the cell and dynamical effects are neglected (Coulomb hole plus screened exchange approximation—COHSEX) [21].

Different methods for the evaluation of ϵ_∞ have been adopted in the past, ranging from self-consistent approach [28] to simplified models or simply tacking the experimental value of the dielectric constant. The hybrid functional built in this way was demonstrated to be able to reproduce the energy gap of a large number of systems. Furthermore it allowed to give a reliable description of the wavefunction properties of defect states in oxides and a good estimate of total and chemical reaction energies [29, 30]. In addition exploiting localized basis set for the evaluation of the exact exchange reduces the computational cost of large supercell calculations.

In the next section we will present the results obtained for the test system of a single vacancy in silicon.

We will show the advantages in using hybrid functionals with respect to semilocal GGA approximation analysing the band structure of the system. Although the one-electron Kohn-Sham eigenvalues can not be related to the experimental electronic levels in the gap, they can give a qualitative idea of the improvement that can be obtained using hybrid functionals. A more accurate evaluation of the defect states in the gap, which is beyond the aim of this report, requires the calculation of the charge transition levels obtained as difference between the formation energy of the defect in two different charge states [31, 32].

2.3 Test Case: Single Vacancy in Silicon

The silicon vacancy is one of the most well-studied defects in semiconductors, and is often treated as a test case for theoretical and computational methods. The vacancy and related defect complexes in silicon have been studied experimentally using several different techniques (for a review, see [33]). A canonical description of this defect centre is given by the Watkins model of the silicon vacancy which was deduced by EPR and from considerations based on the Jahn-Teller theorem [34]. The model allowed to explain the symmetry information given by the experiments, showing that the four dangling bonds of the atoms surrounding the vacancy should combine in a s state which lie in the bulk valence band and three p -states that can be degenerate only in the v^{2+} charge state. Differently, in the neutral vacancy the degeneracy is lifted due to the Jahn-Teller distortion that lowers the symmetry from T_d to D_{2d} . A consequence of the lattice relaxation for different charge states is the unexpected ordering of the stable charge transition levels. Indeed this defect exhibits a negative-effective-U [31], by which the stable charge state changes directly from v^{2+} to v^0 as the electronic chemical potential is raised; the v^{1+} state is therefore only metastable.

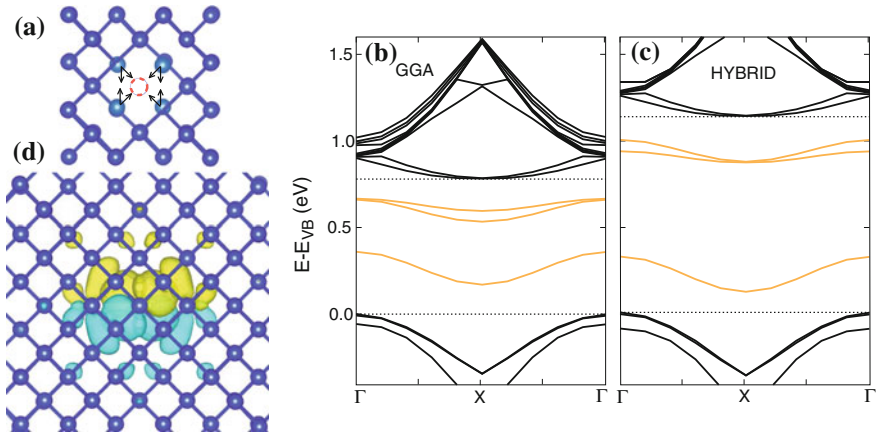


Fig. 2.1 **a** Optimized geometry for the single vacancy in silicon. Light blue atoms correspond to the nearest neighbours of the vacancy (white dot). Arrows indicate their shift upon relaxation. In panel **b** and **c** the band structure of the silicon vacancy from GGA and screened exchange hybrid calculations, respectively. Orange (light grey) bands are defect-related while dashed lines indicates the top of the valence band and the bottom of the conduction band. Panel **d** shows the spin-density of the single vacancy in silicon obtained through the exchange hybrid calculation

Theoretical electronic properties of silicon vacancy were widely studied in the past, also considering the finite size convergence properties [35–37].

These calculations were performed using local LDA and semi-local GGA approximation for the exchange and correlation term. Differently, we compare the results obtained in GGA with Perdew-Burke-Erzenov (PBE) functional with those obtained with the screened exchange hybrid functional.

Our calculation is performed using the all-electron DFT method with a localized atom-centred basis set implemented in the CRYSTAL14 code [38]. The exact exchange fraction is set equal to 0.084, as derived from the experimental value of the static dielectric constant (11.9) [39].

A $3 \times 3 \times 3$ supercell is considered and all the 215 atoms (216 without the vacancy) have been relaxed. The geometry optimization reproduces the expected local arrangement with nearest neighbours of the vacancy forming pairs, as can be observed by the structure reported in Fig. 2.1a. We consider only the singlet spin configuration ($S = 0$) which is known to be the ground state one. Differently, for other kind of defect in silicon a stability test as a function of the chosen spin ground state should be addressed. The basis set for silicon is the Si 88-31 G(*) reported in [40].

Figure 2.1b, c report the band structure obtained with GGA and with the screened exchange hybrid functional, respectively. The first relevant difference between the two results is relative to the energy gap which is underestimated in GGA (0.78 eV) while reproduces almost exactly the experimental value (1.14 eV) in the hybrid

functional calculation showing that the screened exchange is suitable to reproduce the silicon properties.

In both cases three defect states are found within the energy gap, corresponding to the three p-states whose degeneracy is lifted due to the atomic distortion around the defect. Only the lowest of these state is occupied by two electrons with opposite spin. The s-like state lies in the valence band below the energy range considered. Even the defect-related bands display a finite dispersion which is due to the interaction between periodic replica of the vacancy (physically corresponding to a very large defect concentration). Notably, the bands related to the defect state are less dispersing in the calculation with the hybrid potential reflecting a reduction of delocalization of the charge density. This is confirmed also by the behaviour of the spin density (Fig. 2.1d). This quantity has a relevant contribution on the nearest neighbours of the vacancy while it rapidly decays far away from the defect. The total spin moment is equal to zero because different atoms in the cell have opposite contribution leading to a total $S = 0$ state. Comparing the spin density in the hybrid and GGA calculation (not shown) we found that the latter is smaller than the former on the nearest neighbours of the defect and it is larger on the cell boundary, confirming a larger delocalization of the electrons. In the band structure of Fig. 2.1c the energy position of the defect bands in the gap is in quite good agreement with the experimental defect levels [41], being at -0.25 eV and $+0.13$ eV from the conduction band minimum and valence band maximum, respectively. These values are only qualitative and their variation due to the transition to a different charge state should be considered for a more reliable estimate, as discussed above. It is worth noting that the hybrid potential has not simply the effect of a scissor operator but it allows a reorganization of the atomic levels that display different properties with respect to GGA (dispersion, energy alignment with the band edges).

2.4 Conclusion

In this report we have tackled the problem of an accurate description of atomic defects in semiconductors within the Density Functional Theory approach. The final purpose is to characterize the local atomic structure, the defect levels in the band gap and the distribution of the related excess charge in the neighbourhood of the defect. These informations would be a useful tool for the interpretation of experimental measurements on silicon samples with implanted dopant atoms. In order to address these points it is necessary to go beyond DFT-Kohn-Sham local or semilocal functionals such as LDA or GGA, which, besides underestimating the gap, are known to overestimate the delocalization of the excess charge in the host material. We assessed a method based on hybrid functionals which is less demanding than high-level approaches (as GW) and we show its successful applicability in the simple case of a silicon vacancy.

References

1. M. Fuechsle, J.A. Miwa, S. Mahapatra, H. Ryu, S. Lee, O. Warschkow, L.C.L. Hollenberg, G. Klimeck, M.Y. Simmons, A single atom transistor. *Nat. Nanotechnol.* **7**, 242 (2012)
2. D. Moraru, A. Udhiarto, M. Anwar, R. Nowak, R. Jablonski, E. Hamid, J.C. Tarido, T. Mizuno, M. Tabe, Atom devices based on single dopants in silicon nanostructures. *Nanoscale Res. Lett.* **6**, 479 (2011)
3. M. Pierre, R. Wacquez, X. Jehl, M. Sanquer, M. Vinet, O. Cueto, Single-donor ionization energies in a nanoscale CMOS channel. *Nat. Nanotechnol.* **5**, 133 (2010)
4. E. Prati, M. Hori, F. Guagliardo, G. Ferrari, T. Shinada, Anderson-Mott transition in arrays of a few dopant atoms in a silicon transistor. *Nat. Nanotech.* **7**, 443 (2012)
5. E. Prati, K. Kumagai, M. Hori, T. Shinada, Band transport across a chain of dopant sites in silicon over micron distances and high temperatures. *Sci. Rep.* **6**, 19704 (2016)
6. Y.R. Suprun-Belevich, L. Palmethofer, Deep defect levels and mechanical strain in Ge⁺-implanted silicon. *Nucl. Instrum. Methods Phys. Res. B* **96**, 245 (1996)
7. G.D. Watkins, *Defects and Their Structure in Non-metallic Solids* (B. Henderson and A. E. Hughes (Plenum), New York, 1976)
8. W. Kohn, L.J. Sham, Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, A1133 (1965)
9. J. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865 (1996)
10. S. Pagliara, S. Tognolini, L. Bignardi, G. Galimberti, S. Achilli, M.I. Trioni, W.F. Van Dorp, V. Ocelk, P. Rudolf, F. Parmigiani, Nature of the surface states at the single-layer graphene/Cu(111) and graphene/polycrystalline-Cu interfaces. *Phys. Rev. B* **91**, 195440 (2015)
11. S. Achilli, S. Caravati, M.I. Trioni, Ab initio electronic and magnetic properties of 1 ML Fe/Cu(001). *J. Phys. Condens. Matter* **19**, 305021 (2007)
12. S. Achilli, M.I. Trioni, G.P. Brivio, Detailed features of the surface electronic states of K/Cu(111) by density functional theory. *Phys. Rev. B* **81**, 165444 (2010)
13. E. Del Castillo, S. Achilli, F. Cargnoni, D. Ceresoli, R. Soave, M.I. Trioni, Spin-filtering in graphene junctions with Ti and Co adsorbates. *Chem. Phys.* **478**, 1 (2016)
14. P. Bonardi, S. Achilli, G.F. Tantardini, R. Martinazzo, Electron transport in carbon wires in contact with Ag electrodes: a detailed first principles investigation. *Phys. Chem. Chem. Phys.* **17**, 18413 (2015)
15. S. Achilli, M.I. Trioni, E.V. Chulkov, Self-consistent approach for spectral properties of single alkali adatoms on Cu(111). *Phys. Rev. B* **85**, 045408 (2012)
16. M.I. Trioni, G. Fratesi, S. Achilli, G. P. Brivio, Dynamics of electron distributions probed by helium scattering. *J. Phys. Condens. Matter* **21**, 264003 (2009)
17. E. Del Castillo, F. Cargnoni, S. Achilli, G.F. Tantardini, M.I. Trioni, Spin asymmetric band gap opening in graphene by Fe adsorption. *Surf. Sci.* **634**, 62 (2015)
18. J.P. Perdew, A. Zunger, Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B* **23**, 5048 (1981)
19. J.P. Perdew, R.G. Parr, M. Levy, J.L. Balduz Jr., Density-functional theory for fractional particle number: derivative discontinuities of the energy. *Phys. Rev. Lett.* **49**, 1691 (1982)
20. G. Onida, L. Reining, A. Rubio, Electronic excitations: density-functional versus many-body Greens-function approaches. *Rev. Mod. Phys.* **74**, 601 (2002)
21. L. Hedin, New method for calculating the one-particle green's function with application to the electron-gas problem. *Phys. Rev.* **139**, A796 (1965)
22. P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G.L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. Fabris, G. Fratesi, S. de Gironcoli, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, R.M. Wentzcovitch, Quantum ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys. Condens. Matter* **21**, 395502 (2009)

23. P. Umari, G. Stenuit, S. Baroni, Optimal representation of the polarization propagator for large-scale GW calculations. *Phys. Rev. B* **79**, 201104(R) (2009)
24. P. Umari, G. Stenuit, S. Baroni, GW quasiparticle spectra from occupied states only. *Phys. Rev. B* **81**, 115104 (2010)
25. J.P. Perdew, M. Ernzerhof, K. Burke, Rationale for mixing exact exchange with density functional approximations. *J. Chem. Phys.* **105**, 9982 (1996)
26. C. Lee, W. Yang, R.J. Parr, Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **37**, 785 (1988)
27. J. Heyd, G.E. Scuseria, M. Ernzerhof, Hybrid functionals based on a screened Coulomb potential. *J. Chem. Phys.* **118**, 8207 (2003)
28. J.H. Skone, M. Govoni, G. Galli, Self-consistent hybrid functional for condensed systems. *Phys. Rev. B* **89**, 195112 (2014)
29. M. Gerosa, C.E. Bottani, L. Caramella, G. Onida, C. Di Valentin, G. Pacchioni, Electronic structure and phase stability of oxide semiconductors: performance of dielectric-dependent hybrid functional DFT, benchmarked against GW band structure calculations and experiments. *Phys. Rev. B* **91**, 155201 (2015)
30. M. Gerosa, C. Di Valentin, C.E. Bottani, G. Onida, G., Pacchioni, Hole localization in Al-doped quartz SiO₂ within ab initio hybrid-functional DFT. *J. Chem. Phys.* **143**, 111103 (2015)
31. G.A. Baraff, O. Kane, M. Schlüter, Theory of the silicon vacancy: an Anderson negative-U system. *Phys. Rev. B* **21**, 5662 (1980)
32. F. Corsetti, A.A. Mostofi, System-size convergence of point defect properties: the case of the silicon vacancy. *Phys. Rev. B* **84**, 035209 (2011)
33. G.D. Watkins, *Deep Centres in Semiconductors* (S. T. Pantelides Gordon and Breach, New York, 1986)
34. H.A. Jahn, E. Teller, Stability of polyatomic molecules in degenerate electronic states. I: orbital degeneracy. *Proc. R. Soc. Lond. A* **161**, 220 (1937)
35. M.J. Puska, S. Pöykkö, M. Pesola, R.M. Nieminen, Convergence of supercell calculations for point defects in semiconductors: vacancy in silicon. *Phys. Rev. B* **58**, 1318 (1998)
36. A.F. Wright, Density-functional-theory calculations for the silicon vacancy. *Phys. Rev. B* **74**, 165116 (2006)
37. M.I.J. Probert, M.C. Payne, Improving the convergence of defect calculations in supercells: an ab initio study of the neutral silicon vacancy. *Phys. Rev. B* **67**, 075204 (2003)
38. R. Dovesi, R. Orlando, A. Erba, C. Zicovich-Wilson, B. Civalleri, S. Casassa, L. Maschio, M. Ferrabone, M. De La Pierre, P. D'Arco, Y. Noel, M. Causa, M. Rerat, B. Kirtman, CRYSTAL14: a program for the ab initio investigation of crystalline solids. *Int. J. Quantum Chem.* **114**, 1287 (2014)
39. P.Y. Yu, M. Cardona, *Fundamentals of Semiconductors* (Springer, Berlin, 2001)
40. R. Nada, J.B. Nicholas, M.I. McCarthy, A.C. Hess, Basis sets for ab initio periodic Hartree-Fock studies of zeolite/adsorbate interactions: He, Ne, and Ar in silica sodalite. *Int. J. Quantum Chem.* **60**, 809 (1996)
41. V.V. Lukjanitsa, Energy levels of vacancies and interstitial atoms in the band gap of silicon. *Semiconductors* **37**, 404 (2003)

Chapter 3

Gas and Dust Dynamics During Planet Formation in HL Tau



Giulia Ballabio, Giuseppe Lodato and Giovanni Dipierro

Abstract The disc planet interaction and planetary formation in accretion discs have become a topic of renewed interest in the last decade, encouraged by the discovery of extra-solar planets, and in particular of hot Jupiters. Recent long-baseline ALMA (Atacama Large Millimeter Array) observations revealed a striking pattern of bright and dark rings in the protoplanetary disc surrounding the young star HL Tau. Our group has provided one of the earliest explanations of this system in terms of the interaction of the disc with three newly born planets (Dipierro et al, MNRAS 453, L73–L77, 2015) [6]. The aim of this study is to develop a new model for gas and dust distribution within this disc. This work is focused on identifying the mass of the planets able to carve gaps in HL Tau, in the three main ring-like structures observed by ALMA.

3.1 Introduction

It is a well known fact that stars in the solar neighbourhood are formed from the gravitationally induced collapse of cold molecular gas. As a results of this collapse, a gaseous and dusty protoplanetary disc orbits around the young protostar for the first few million years of its evolution [10]. It is generally assumed that these protoplanetary discs are the birthplaces of planets. The disc planet interaction and planetary formation in accretion discs have become a topic of renewed interest in the last decade, encouraged by the discovery of extra-solar planets, and in particular of hot Jupiters. Recent technical advances, such as the completion of the Atacama Large Millimeter/submillimeter Array (ALMA), are offering us an unprecedented

G. Ballabio (✉) · G. Dipierro

Department of Physics and Astrophysics, University of Leicester, Leicester, UK
e-mail: gb258@leicester.ac.uk

G. Ballabio · G. Lodato

Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy

© Springer Nature Switzerland AG 2018

P. F. Bortignon et al. (eds.), *Toward a Science Campus in Milan*,
https://doi.org/10.1007/978-3-030-01629-6_3

opportunity to gain insights on protoplanetary discs and planet formation process. Observations indeed are revealing a wide range of disc substructures with an impressive detail (on a scale of a few au), that might be connected to the presence of forming protoplanets. One of the most popular and surprising discoveries of the latest decade is the disc around the young star HL Tau. This was the first resolved observation made by ALMA in full regime, with the most extended configuration and highest sensitivity.

Of particular interest for us are detections of gaps and ring-like features. Despite the large number of mechanisms that could potentially induce gaps opening in protoplanetary discs, the planetary hypothesis is the widely invoked mechanism to explain their formation. It is known that an embedded protoplanet interacts gravitationally with the surrounding material, leading to a local exchange of angular momentum that induce perturbations both in the disc structure and in the planets orbit (see [8, 9, 13, 14]). In the limit of a very low mass perturber, the tidal interaction does not affect the gas structure, providing a uniform flux of angular momentum. For a sufficiently high mass planet instead, the tidal torques on the surrounding gas overcome the pressure gradient and the viscous spreading. As a result, the planet can alter the gas density in its vicinity, pushing material away from its orbit and leading to the formation of a gap [2]. Dust, by contrast, is a pressureless and inviscid fluid: the formation of gaps in this case is due to the competition between the tidal torque and the aerodynamic drag torque. As a consequence, the criterion for gap opening in dusty disc depends upon the grain size of particles as well as planetary mass (see [5, 18]). One of the main beliefs is that gap formation in gaseous disc is not a necessary condition for opening a gap in the dust component.

3.2 From Observations to the Theoretical Model

The key to figure out the nature of these substructures is based on the investigation of the spatial distribution of the dust and the gas phases within protoplanetary discs. Although the mass of these discs is dominated by gas, the solid component plays a crucial role in the dynamic evolution and thermal structure of the system, since it dominates the opacity budget [24]. For this reason, observations of light scattered on dust grains and their direct thermal emission remain the key to detect discs and to characterize their structure. The structure of the dust and gas spatial distribution has to be correctly computed in order to figure out which mechanisms is responsible for such substructures. By identifying these structures and comparing them with theoretical models, it is possible to provide predictive links to the local gas and dust conditions. This work is focused at reproducing the observational features of discs hosting planets and consists in a follow-up of a previous work (see [6]).

3.2.1 Our Goal: HL Tau

Recent long-baseline observations of ALMA revealed a striking pattern of bright and dark rings in the protoplanetary disc surrounding the young star HL Tau [1]. HL Tau is a T Tauri star evolving from SED Class I to Class II, located in the Taurus molecular cloud, at a distance of 140 pc. Its disc has been widely studied in the sub-millimeter and millimeter continuum imaging of interferometers, as well as molecular line observations. One of the main purposes of this study relies in building a new model for HL Tau, in order to set up suitable initial conditions for hydrodynamics simulations (see the next section).

A good point to start is considering the model of [11], modified taking into account additional results presented in the most recent papers of [4, 12, 19]. The disc model considered predicts a power law radial density distribution, tapered by an exponential function, which in the case of a thin disc has the form

$$\Sigma_{\text{d}}(R) = \Sigma_{\text{d},0} \left(\frac{R}{R_{\text{c}}} \right)^{-p_{\text{d}}} \exp \left[- \left(\frac{R}{R_{\text{c}}} \right)^{2-p_{\text{d}}} \right]. \quad (3.1)$$

It is crucial to determine the parameters of the model: the p -index, the characteristic radius R_{c} and the normalisation $\Sigma_{\text{d},0}$. At this point we have to dwell on the fact that we are dealing with quantities evaluated on the mid-plane, where planets are expected to form. Nevertheless, the emission from the inner regions in HL Tau turned out to be optically thick at ALMA wavelengths, preventing the derivation of surface density profiles and grain size distributions. For this reason, we consider the more sensitive observations of HL Tau obtained through the Very Large Array (VLA). Carrasco-Gonzalez [4] observed HL Tau at a wavelength of 7.0 mm, with a spatial resolution comparable to the ALMA images. At this long wavelength the dust emission from HL Tau is optically thin, allowing a comprehensive study of the inner disc.

In order to build a new model for HL Tau, we therefore carry out a fit considering the observations from [4] for the inner part of the disc and the data from [19] for the outer part of the disc. The fit has three parameters and the fitting function has the form of the dust surface density profile of Sect. 3.1. The result of our fit is plotted in the left panel of Fig. 3.1. The significant parameters of the fit are $p = 1.0$ and $R_{\text{c}} = 29.4$ au. The normalisation coefficient is consistent with the estimate taken for the dust mass.

We now derive the radial profile for the gas surface density. As already said, in protoplanetary discs the gaseous component is dominant, for a value of 100 by mass. Unfortunately, since it is rather difficult to be observed, we do not have so much observational restriction as in the case of the dust profile. However, the gas surface density is involved in determining the Stokes number radial profile on the mid-plane. We thus recall the definition of the Stokes number

$$\text{St} = \frac{\pi r \rho_{\text{m}}}{2 \Sigma_{\text{g}}}. \quad (3.2)$$

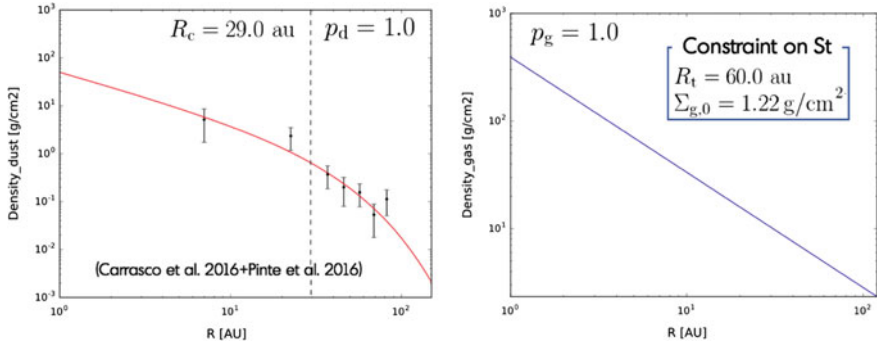


Fig. 3.1 Surface density profile of dust (left) and gas (right) components. In the left panel, black dots refer to observational data and the fitting red line is a power law tapered by an exponential function.

Table 3.1 Set of parameters we used as initial conditions for our SPH simulations

Initial conditions

$M_{\text{star}} (M_{\odot})$	1.7	$M_{\text{gas}} (M_{\odot})$	0.006
$R_{\text{in}} (\text{au})$	1.0	p_{g}	1.0
$R_{\text{out}} (\text{au})$	120.0	$R_{\text{t}} (\text{au})$	60.0
α	$5 \cdot 10^{-3}$	$M_{\text{dust}} (M_{\odot})$	$5 \cdot 10^{-4}$
$(H/R)_0$	0.05	p_{d}	1.0
q	0.215	$R_{\text{c}} (\text{au})$	29.0
$\rho_{\text{m}} (\text{g/cm}^3)$	3	$r (\text{mm})$	1.3

and we require that grains with millimeter size must reach a Stokes number near to unity, at the planet location: we therefore set $\text{St} = 0.5$ at $R = 60$ au, for 1.3 mm grains. This assumption allow us to derive the radial gas density profile, shown in the right panel of Fig. 3.1.

Up to now we have only acquired a new analytical model derived from the most recent observations, in order to describe the structure of a disc with both a gaseous and a dusty counterpart. In Table 3.1 we summarize all the parameters obtained in our analysis. These are the initial conditions we need to set in order to simulate the evolution of an accretion protoplanetary disc.

3.3 Smoothed Particle Hydrodynamics Simulations

In an astrophysical context, two types of numerical simulations are generally adopted: grid-based codes [18] and particle-based smoothed particle hydrodynamics (SPH) codes [16, 17]. SPH is a numerical approach for non-axisymmetric problems in

astrophysics involving fluid dynamics. It is a Lagrangian method, which allows the conservation of linear and angular momentum and of other important physical invariants.

In order to study the planetary formation within protoplanetary discs, we employ the PHANTOM SPH code (see [15, 20, 21, 23]), developed by Daniel Price (and now public). PHANTOM is a fast, parallel, modular and low-memory smoothed particle hydrodynamics and magnetohydrodynamics code developed over the last decade for astrophysical applications in three dimensions. In PHANTOM, hydrodynamics fluid equations are computed consistently, taking into account gas viscosity, gravity and drag forces which are involved in the planetary formation process.

In this work we perform global 3D simulations of both gas and dust within a protoplanetary disc, hosting one or more embedded planets. After several test with only one embedded planet, we deal with the real system we want to investigate: a protoplanetary disc hosting three planets as those observed in HL Tau. Through ALMA observations, [1] derived the position of the gaps at 13.2, 32.3 and 68.8 au. These locations corresponds to resonances 12 : 3 : 1 for the orbital frequencies. Whereas, starting from the estimate given by [6], we carry out a fine-tuning on the values of planet masses. In Fig. 3.2 we show the results obtained by the most remarkable of our simulations. We adopt a system which comprises a central star with mass $M_{\text{star}} = 1.7 M_{\odot}$ and an accretion radius of $R_{\text{acc}} = 1$ au. It is surrounded by a disc of $5 \cdot 10^5$ SPH particles which extends from $R_{\text{in}} = 1$ au to $R_{\text{out}} = 120$ au. We assume a radial isothermal equation of state, such that the gas temperature power law is defined by $q = 0.35$. The aspect ratio at the inner radius is set at $(H/R)_0 = 0.02$. The gaseous disc weighs $0.05 M_{\odot}$, whose distribution is governed by a power law profile with $p_g = 1.0$. The dust is allocated according to its radial power law profile

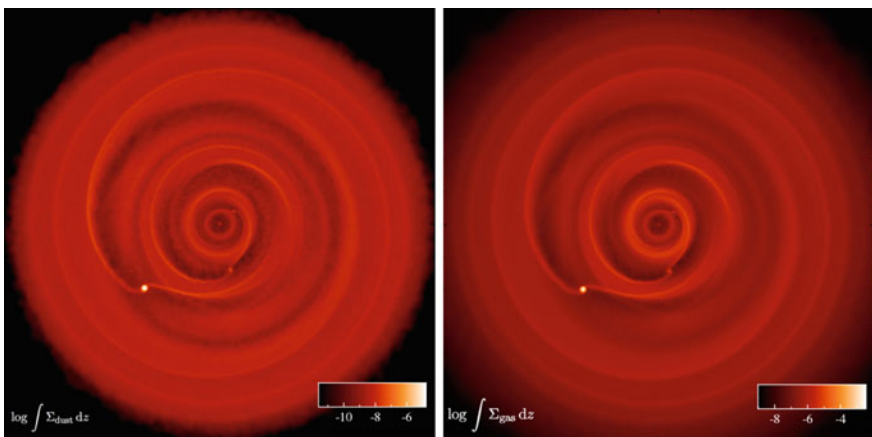


Fig. 3.2 Rendered images of a disc with three embedded planets at orbital radii of 13.2, 32.3 and 68.8 au. The masses estimated for each planet are 0.5, 0.7 and 1.0 MJ. After 11 orbits of the outer planet, shallow gaps can be observed in the dust phase (left panel). The gas develops similar but smoother structures (right panel)

with index $p_d = 0.8$. Finally, we assume a grain density of 0.5 g/cm^3 . We therefore place three planets at given radii, with masses equal to $0.5, 0.7$ and $1.0 M_J$.

3.3.1 *Traces of Gaps Formation*

Figure 3.2 illustrates rendered images of both the dust (left) and the gas (right) surface density from the fiducial disc model. It is evident the presence of forming shallow gaps shaped by planets, both in the dust and in the gas. After 45 orbits of the second planet, the structure of the innermost gap is totally vanished, and material is accreted onto the star. The dust phase is more affected by this lost of material occurring in the central region. This cause a broadening of the edges of the second gap and a depletion of dust grains. We can notice instead that this gap is more delineated in the gas phase. The formation of such structures as that illustrated in Fig. 3.2 occurs after ~ 11 orbits of the outer planet (32 for the second and 130 for the inner planet).

In conclusion, we can not certain confirm the presence of clear cavities as such derived by simulations performed in [6]. At the moment we are not able to reproduce the axisymmetric structures observed in HL Tau. The gaps in our simulations are not so clean and well defined. The main reason lies behind the violation of dust mass conservation. Indeed, nevertheless we performed a fine-tuning on almost all the parameters of our model, the problem is still present: we only reduce its effects. All our simulations reveal an amount of dust mass over time, largely attending at outer radii and in the upper layer of the disc. A possible solution lies in trying to fix the bug directly in the code: this is the next step of our work.

3.4 **Simulating Dust in SPH**

Within discs the dynamics of small dust grains is typically well coupled to that of the gas. For larger grains, however, the dust and gas dynamics can be decoupled. A dust-gas mixture is usually treated using a continuous two fluid description, and a large class of numerical solvers have been developed. In SPH, to compute the evolution of this mixture there are two possible methods:

- the two fluid implementation, where the dust and gas are treated as two separate fluids coupled by a drag term;
- the one fluid implementation, in which the dust is treated as part of the mixture, with an evolution equation for the dust fraction.

Approaches for modelling the dynamics of dust grains that are decoupled from the motions of the gas are often distinguished by whether they use a single or two fluid method. The strong coupling between the gas and the dust phases has to be resolved,

in order to perform simulations with the two fluid implementation. The condition to resolve differential motion between the two fluids is

$$h < c_s t_s. \quad (3.3)$$

The more the stopping time is small, the more the scale to be resolved is tiny. Thus we should increase the spatial resolution. To reach higher resolutions is then necessary to increase the number of particles, meaning a bigger computational cost. Unfortunately, small stopping time are required, due to the small values of the Stokes number, especially in the inner part of the disc. If we want to resolve the gas and the dust dynamics at smaller radii, we are forced to make use of the single fluid algorithm.

3.4.1 Violation of Dust Mass

During this work, we discovered that the one fluid hydrodynamics simulations are affected by a strong violation of dust mass conservation. This unphysical effect is clearly visible at large radii in the upper layers of the disc and it is important to find the cause. In some cases, as the one presented in this study, violation of the dust mass conservation equation

$$\frac{d\epsilon}{dt} = -\frac{1}{\rho} \nabla \cdot [\epsilon(1 - \epsilon)\rho \Delta \mathbf{v}]. \quad (3.4)$$

can occur if the dust fraction—the ratio between the dust density and the total density—reaches critical values. Whenever $\epsilon \rightarrow 0$ and/or $\epsilon \rightarrow 1$ the evolution equation might diverges.

In this section we thus exhibit a new implementation for the dust fraction similar to that used by [22], but defining different variables. The constraint is $0 < \epsilon < 1$. The idea is to redefine the dust fraction as a function whose co-domain is defined on $[0, 1]$, thus preventing ϵ to overflow outside this range. It is therefore of great importance to choose the most suitable function. As we experienced, this is a hard issue, since it is not enough to find a function with this peculiarity. A more difficult problem arises when we compute the evolution equation for the new variables (as we will show below). The decisive solution would be to find a clever way to express the derivative, as illustrated in [22].

We therefore derived new implementations of the variable describing the dust fraction in the code, showing here just one example. We started considering the function $\epsilon = \sin^2 \theta$, where θ is the new variable. In order to express θ as a function of ϵ , the function $\sin^2 \theta$ must be invertible. We thus limit its domain to $\theta \in [0, \pi/2]$, setting some restraining conditions in the code, and we obtain

$$\theta = \frac{1}{2} \arccos(1 - 2\epsilon). \quad (3.5)$$

The time derivative is therefore

$$\frac{d\theta}{dt} = -\frac{1}{2} \frac{1}{\rho\sqrt{\epsilon(1-\epsilon)}} \nabla \cdot (\epsilon t_s \nabla P). \quad (3.6)$$

It is obvious that this expression diverges if either ϵ or $(1 - \epsilon)$ are zero.

3.4.2 1D NDSPMHD Simulations

To verify if functions we choose could be a better solution for the implementation of the dust fraction, we turn to 1D SPH simulations. We therefore use the NDSPMHD code, which is thought to be a code for algorithmic experimentations. It includes full implementations of 1D, 2D and 3D hydrodynamics and magnetohydrodynamics, as described in [20]. We therefore performed 1D SPH simulations reproducing the dust settling, in order to test these new implementations.

3.4.3 Dust_settle Test

This is a 1D test of dust settling in a vertical section of a protoplanetary disc. It is one of the test problems already implemented in the code. We set up 10^5 SPH particles in a gas disc with total mass of $0.01 M_\odot$ surrounding a 1 Solar mass star. The initial dust-to-gas ratio is 0.01, so the dust only slightly affects the gas motion. The disc extends from 10 to 400 au. Both gas and dust phase are placed using a surface density profile such that the power law index is fixed, $\Sigma(R) \propto R^{-1}$. The radial profile of the gas temperature is taken to be $T(R) \propto R^{-0.6}$, with $H/R = 0.05$ at 100 au. Since a uniform grain size of 1 mm is used (with $\rho_m = 3 \text{ g/cm}^3$), we can employ the Epstein formulation for the drag force. Results of our 1D simulations of dust settling are illustrated in Fig. 3.3, which shows the evolution of dust fraction as a function of the disc scale height z . Panel (a) clearly shows that the test fails for the \sin^2 function, and dust settling does not take place. In the other case dust settles towards the mid-plane until it reaches a value for $H/R \approx 0.05$. This could be the best choice for our new implementation.

Nonetheless, one of the latest implementations appears to be promising to resolve the numerical issue. We define the dust fraction as

$$\epsilon = \frac{s^2/\rho}{1 + s^2/\rho} = \frac{s^2}{\rho + s^2} \quad \text{such that} \quad s = \sqrt{\frac{\rho\epsilon}{1-\epsilon}}. \quad (3.7)$$

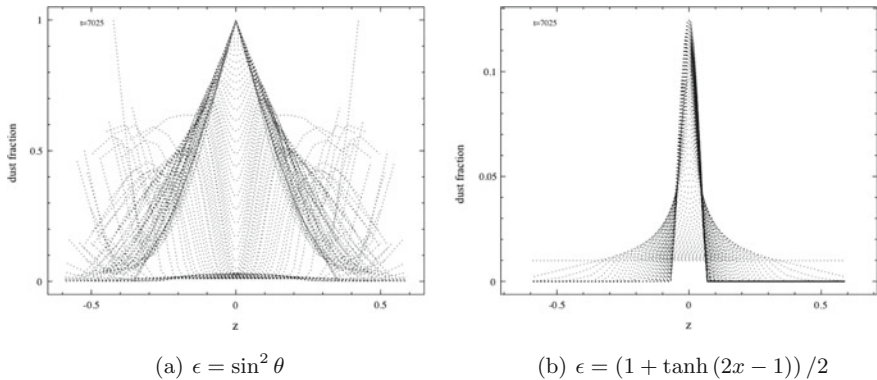


Fig. 3.3 1D simulations of dust settling performed with NDSPMHD. Each panel shows the evolution of dust fraction for the derived functions, after 50 orbits. The horizontal axis is the disc scale height z . The test fails in the case shown in panel (a) whereas in the other case dust settles towards the mid-plane (b)

We calculate the time derivative obtaining

$$\frac{ds}{dt} = -\frac{1}{2} \frac{1}{(1-\epsilon)^2} \nabla \cdot \left(\frac{s(1-\epsilon)}{\rho} t_s \nabla P \right) - \frac{t_s}{2\rho} \frac{1}{(1-\epsilon)} \nabla P \cdot \nabla s - \frac{s}{2} \nabla \cdot \mathbf{v}. \quad (3.8)$$

It can be notice that there is still a problem with the evolution equation if $\epsilon \rightarrow 1$, i.e. for $\rho_g \rightarrow 0$. However, this formulation has been implemented and tested in the code and seems to provide interesting and definitive results [3].

3.5 RADMC-3D and CASA

In addition, Monte Carlo radiative transfer simulations have been performed using the RADMC-3D code [7] in order to compute the expected emission of the disc models for HL Tau. RADMC-3D is a software package for astrophysical radiative transfer calculations in arbitrary 1D, 2D or 3D geometries. It is mainly written for continuum radiative transfer in dusty media, but also includes modules for gas line transfer and gas continuum transfer. Starting from the disc models obtained from hydrodynamics simulations, we translate them into model observations. We focus on ALMA band 6, which corresponds to a continuum emission at 233 GHz ($\lambda = 1287 \mu\text{m}$). The source is located at 140 pc, in the position of HL Tau, adopting the disc inclination ($\theta \sim 47^\circ$) and position angle (PA $\sim 138^\circ$) given by [1]. The dust distribution is used directly

from our simulation data rather than being prescribed. The dust model consists of spherical silicate grains with optical constant for magnesium-iron grains taken from the Jena database.

We also simulate realistic ALMA observations of the adopted theoretical models in order to compare ALMA predictions with real data. The full-resolution submillimeter-millimeter images directly produced by RADMC-3D simulations are used as input sky models to simulate realistic ALMA observations through the Common Astronomy Software Application (CASA) ALMA simulator. This software takes into account the thermal noise from the receivers and the atmosphere and it assumes a perfect calibration of the visibility measurements. A transit duration of 8 hours is used to reach an optimal signal-to-noise ratio. The observation parameters are chosen to match the spatial resolution of the ALMA observations.

3.6 Results

Figure 3.4 shows the results of a radiative transfer simulation performed on the protoplanetary disc described in the previous section. Comparing our results with the real data measured by ALMA, we can notice that there is a good accordance between the intensity scales—they are of the same order of magnitude. This means that our estimate for the dust mass is good. However we are still too far in reproducing the observed cavities in HL Tau, with high precision. We propose this weakness is due to the low value of the Stokes number, currently preventing us from reaching a full and satisfying comprehension of the process of gaps formation in HL Tau. This is just the beginning of a work that will continue in the next few months.

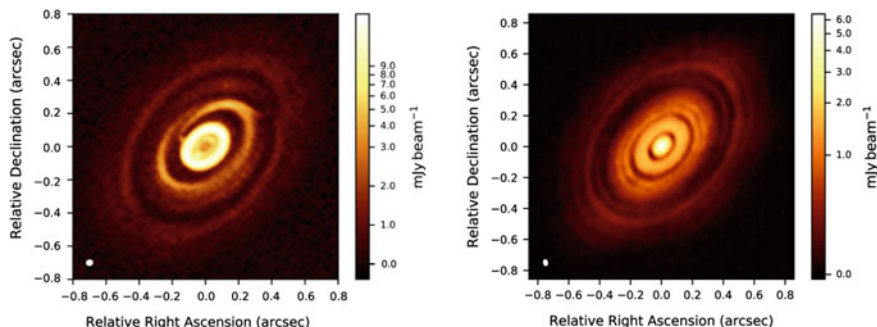


Fig. 3.4 *Three planets*—comparisons between the simulated observations of our disc model at the ALMA Band 6 (left panel) with the ALMA image of HL Tau (right panel). The white colour in the filled ellipse in the lower left corner indicates the size of the half-power contour of the synthesized beam

3.7 Conclusions

In this study we have investigated the spatial distribution of gas and dust within a disc hosting multiple planets and analyzed the observational predictions in order to verify if theoretical models can be interpreted according to recent high resolution observations. In order to reproduce dust rings and gaps structures in HL Tau, we have performed global 3D smoothed hydrodynamics simulations of both gas and dust within a protoplanetary disc sculpted by one or more planets. In addition, the expected emission of the disc model for HL Tau has been computed using Monte Carlo radiative transfer simulations. Starting from the disc models obtained from hydrodynamics simulations, we then translated them into model observations. The full-resolution submillimeter-millimeter images directly produced by RADMC-3D simulations are therefore used as input sky models to simulate realistic ALMA observations (through the software CASA).

However, during this work, we have discovered that the hydrodynamics simulations are affected by a non negligible degree of dust mass, clearly visible at outer radii and in the upper layers of the disc. The big issue is to find what is the cause of this non-conservation. With this aim we have started to manipulate the parameters of the disc model, trying to figure out what influences this unphysical effect. Comparing our results with the real data measured by ALMA, we find a good accordance between the intensity scales—they are of the same order of magnitude. We can conclude that our estimate of the total dust mass is able to reproduce the observed flux emitted by the disc. However we are still too far in reproducing with high precision the observed gaps in HL Tau. We attribute this weakness to the mass spreading that affects hydrodynamics simulations and prevents us to reproduce the axisymmetric observed structures. The gaps in our simulations are not so clean and well defined.

In the next step of this work we started manipulating the code. In numerical simulations a dust-gas mixture is usually treated using a continuous two fluid description. This approach, used in the past, is however too computationally demanding for tightly coupled dust. To compute the evolution of this mixture in SPH, one possible methods is to use a single population of particles to describe the total fluid mass. The one fluid formulation thus treats the dust as part of the mixture, adding an evolution equation for the dust fraction ($\epsilon = \rho_d / \rho_{\text{tot}}$). In this context, the problem of non conservation of dust mass occurs when ϵ reaches the critical values 0 and 1, according to its definition. Therefore we have developed a new implementation for the dust fraction, introducing further modifications in the evolution [3]. Moreover, in a collaboration with Daniel Price (the main code developer), we performed 1D SPH simulations, in order to reproduce the dust settling.

This study can provide a robust path to figure out which is the mechanism behind the planetesimal formations in accretion protoplanetary discs. One of the most intriguing aspects linked to these systems is the characterization of substructures recently detected with unprecedented resolutions. In particular, substructures such as gaps, spirals and horseshoes will be especially important in developing a comprehensive model for the evolution of planetary system.

References

1. A.L.M.A. Partnership, The 2014 ALMA long baseline campaign: first results from high angular resolution observations toward the HL Tau region. *ApJ* **808**, L3 (2015)
2. P.J. Armitage, *Astrophysics of planet formation*. CUP (2010)
3. G. Ballabio, G. Dipierro, G. Veronesi, et al., Enforcing dust mass conservation in 3D simulations of tightly coupled grains with the Phantom SPH code. *MNRAS* **477**, 2766–2771 (2018)
4. C. Carrasco-Gonzalez, T. Henning, The VLA view of the HL Tau disk: disk mass, grain evolution, and early planet formation. *ApJ* **821**, L16 (2016)
5. G. Dipierro, G. Laibe, D. Price, Two mechanisms for dust gap opening in protoplanetary discs. *MNRAS* **459**, L1–L5 (2016)
6. G. Dipierro, D. Price, G. Laibe, K. Hirsh, A. Cerioli, G. Lodato, On planet formation in HL Tau. *MNRAS* **453**, L73–L77 (2015)
7. C.P. Dullemond, A. Juhasz, et al., RADMC-3D: a multi-purpose radiative transfer tool. *Astrophysics Source Code Library* (2012)
8. P. Goldreich, S. Tremaine, The excitation of density waves at the Lindblad and corotation resonances by an external potential. *ApJ* **233**, 857–871 (1979)
9. P. Goldreich, S. Tremaine, Disk-satellite interactions. *ApJ* **241**, 425–441 (1980)
10. L. Hartmann, *Accretion processes in star formation*. CUP (2009)
11. W. Kwon, L.W. Looney, L.G. Mundy, Resolving the circumstellar Disk of HL Tauri at millimeter wavelengths. *ApJ* **741**, 3 (2011)
12. W. Kwon, L.W. Looney, L.G. Mundy, W.J. Welch, Resolving protoplanetary disks at millimeter wavelengths with CARMA. *ApJ* **808**, 102 (2015)
13. D. Lin, J. Papaloizou, Tidal torques on accretion discs in binary systems with extreme mass ratios. *MNRAS* **186**, 799–812 (1979)
14. D. Lin, N.C. Papaloizou, On the tidal interaction between protoplanets and the primordial solar nebula. II—Self-consistent nonlinear interaction. *ApJ* **306**, 395–409 (1986)
15. G. Lodato, D. Price, On the diffusive propagation of warps in thin accretion discs. *MNRAS* **405**, 1212–1226 (2010)
16. J. Monaghan, Smoothed particle hydrodynamics. *ARAA* **30**, 543–574 (1992)
17. J. Monaghan, J. Lattanzio, A refined particle method for astrophysical problems. *AAP* **149**, 135–143 (1985)
18. S. Paardekooper, J. Mellema, Dust flow in gas disks in the presence of embedded planets. *AAP* **453**, 1129–1140 (2006)
19. C. Pinte, W. Dent, et al., Dust and gas in the disk of HL Tauri: surface density, dust settling, and dust-to-gas ratio. *ApJ* **816**, 25 (2016)
20. D.J. Price, Smoothed particle hydrodynamics and magnetohydrodynamics. *JoCP* **231**, 759–794 (2012)
21. D.J. Price, C. Federrath, A comparison between grid and particle methods on the statistics of driven, supersonic, isothermal turbulence. *MNRAS* **406**, 1659–1674 (2010)
22. D.J. Price, G. Laibe, A fast and explicit algorithm for simulating the dynamics of small dust grains with smoothed particle hydrodynamics. *MNRAS* **451**, 813–826 (2015)
23. D.J. Price, J. Wurster, C. Nixon, et al., Phantom: A Smoothed Particle Hydrodynamics and Magnetohydrodynamics Code for Astrophysics. *PASA* **35**, 31 (2018)
24. L. Testi, T. Birnstiel, L. Ricci et al., Dust evolution in protoplanetary disks. *Protostars Planets VI*, 339–361 (2014)

Chapter 4

Quantum Simulation of Non-Markovian Qubit Dynamics by an All-Optical Setup



Claudia Benedetti, Simone Cialdi, Matteo A. C. Rossi, Bassano Vacchini,
Dario Tamascelli, Stefano Olivares and Matteo G. A. Paris

Abstract We address the experimental implementation of a quantum simulator based on an optical setup. Our device can simulate the dynamical evolution of a qubit undergoing a dephasing process. In particular, we focus on the dynamics arising from the interaction with a classical stochastic field. We encode the state of the qubit in the polarization of a single photon, while the realizations of the stochastic evolution affect its spectral components by a programmable spatial-light-modulator. This setup can simulate in one shot the ensemble-averaged dynamics of the dephasing qubit. We experimentally reconstruct the system density matrix and we show how it is possible to move from a Markovian to a non-Markovian quantum map by changing the spectral parameter of the simulated noise.

4.1 Introduction

Simulating the quantum dynamics of a system may be a challenging task. As pointed out by Feynman [1], the problem may be tackled by considering another quantum controllable system that emulates the systems under study.

General-purpose quantum computers would be formidable and flexible tools able to efficiently reproduce the behavior of a large class of quantum systems. But quantum computers have not been built yet. On the other hand, much attention has been recently devoted to the development of *ad-hoc* quantum simulators [2], i.e. special-

C. Benedetti (✉) · S. Cialdi · M. A. C. Rossi · B. Vacchini · D. Tamascelli
S. Olivares · M. G. A. Paris
Quantum Technology Lab, Dipartimento di Fisica “Aldo Pontremoli”,
Università degli Studi di Milano, 20133 Milan, Italy
e-mail: claudia.benedetti@unimi.it
URL: <https://users.unimi.it/aqm>

S. Cialdi · B. Vacchini
INFN, Sezione di Milano, 20133 Milan, Italy

© Springer Nature Switzerland AG 2018
P. F. Bortignon et al. (eds.), *Toward a Science Campus in Milan*,
https://doi.org/10.1007/978-3-030-01629-6_4

purpose devices that are designed to mimic the dynamics of other specific complex quantum systems that are not easily accessible or controllable, such as large quantum systems, with many degrees of freedom. The inherent parallel structure of quantum simulators make them the perfect candidates to solve problems that are intractable on conventional supercomputers.

In this paper, we explain in detail how it is possible to build a quantum simulator using only optical elements. In particular, we describe how to mimic the dynamical evolution of a single qubit noisy channels originating from the interaction with different fluctuating stochastic fields [3, 4]. The simulated dynamics corresponds to effective models for the interaction of qubits with complex classical environments [5, 6]. We encode the state of a qubit in the polarization degree of freedom of a single photon, while the spectral components play the role of the external environment. The different realizations of the stochastic fluctuations are numerically simulated and imposed to the dynamics via a spatial light modulator.

As a proof of principle, we simulate the dynamics of a qubit undergoing a dephasing dynamics; this is done by performing a sample average over a large number of parallel realizations of the stochastic fluctuations. We moreover provide examples of quantum channels with different degrees of non-Markovianity.

The paper is organized as follows: in Sect. 4.2 we introduce the theoretical model of a dephasing map induced by classical noise; in Sect. 4.3 we explain in detail the implementation and functioning of our quantum simulator. Section 4.4 closes the paper with final remarks.

4.2 The Model

We aim to simulate the dynamics of a qubit interacting with a classical fluctuating field $X(t)$, whose evolution is described by a stochastic process. The time-dependent stochastic Hamiltonian of the system is described by:

$$H(t) = \epsilon\sigma_z + X(t)\sigma_x \quad (4.1)$$

where ϵ is the energy splitting of the qubit and σ_z is the Pauli matrix. Since $X(t)$ is a stochastic process, the evolved qubit state is calculated by performing the ensemble average over all possible realizations of the noise, namely

$$\rho(t) = \langle U(t)\rho(0)U^\dagger(t) \rangle_{\{X(t)\}} \quad (4.2)$$

for any initial state $\rho(0)$, the unitary operator $U(t) = \exp[-if H(s)ds]$ represents, instead, the evolution as determined by a particular realization of the process $X(t)$. Without losing generality, we can fix the initial state of the qubit to be the $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ state. The generated dynamics in the interaction picture is calculated from (4.2) and its general form is:

$$\rho(t) = \frac{1}{2} \begin{pmatrix} 1 & \langle e^{-2i\varphi(t)} \rangle \\ \langle e^{2i\varphi(t)} \rangle & 1 \end{pmatrix} \quad (4.3)$$

where $\langle e^{-2i\varphi(t)} \rangle$ is the decoherence factor and $\varphi(t) = \int_0^t X(s)ds$. The decoherence factor can be calculated only once the master equation of the stochastic process $X(t)$ is known. To this aim we consider two stochastic processes, the random telegraph noise (RTN) and the Ornstein-Uhlenbeck (OU) process [7–9]. The generated dynamics is a dephasing map, which preserves the populations in the computational basis and only affects the off-diagonal elements through the decoherence factor.

The RTN is a Markovian Gaussian stochastic process. In this case $X(t) \in \{+1, -1\}$, namely it switches between two values, with a fixed switching rate γ . It is characterized by an exponential autocorrelation function

$$C(t - t_0) = e^{-2\gamma|t-t_0|}. \quad (4.4)$$

If we discretize time and consider a small time interval δt , then the probability for the RTN to flip its value within δt is given by $\delta P = 1 - e^{-\gamma\delta t}$.

On the contrary, the OU process is a Gaussian process, described by the stochastic equation

$$X(t + \delta t) = (1 - 2\gamma\delta t)X(t) + 2\sqrt{\gamma}\delta W(t), \quad (4.5)$$

with $dW(t)$ a Wiener process with zero mean and standard deviation $\sigma = \sqrt{\gamma}$. The OU process has the same autocorrelation function as the RTN, even though the statistics of the two processes are different. For both RTN and OU the decoherence factor is a real quantity, and takes the expressions:

$$\langle e^{-2i\varphi(t)} \rangle_{\text{RTN}} = e^{-\gamma t} \left(\cosh \delta t + \frac{\gamma}{\delta} \sinh \delta t \right) \quad \text{with } \delta = \sqrt{\gamma^2 - 4}, \quad (4.6)$$

$$\langle e^{-2i\varphi(t)} \rangle_{\text{OU}} = e^{-2\beta(t)} \quad \text{with } \beta(t) = \frac{1}{\gamma^2} (\gamma t + e^{-\gamma t} - 1). \quad (4.7)$$

Given a certain quantum open dynamics, we can ask if memory effects play a role in the characteristics of such map. As classical processes can be classified into Markovian and non-Markovian (NM), various attempts to generalize such memory effects also in quantum dynamics have been made [10–12]. However a unique definition does not exist. One proposal consists in evaluating the non-Markovianity of the dynamical map as the amount of information backflow into the system from the environment [13]. This is calculated by analyzing the revivals of the trace distance $C(t) = \frac{1}{2} \text{Tr}|\rho_1(t) - \rho_2(t)|$ between a pair of initial states of the system. Whenever the trace distance has revivals in time, the dynamics is called non-Markovian, while a monotonic behavior of the trace distance is associated to a Markovian evolution. The difficulty in calculating such NM lies in the optimization over all possible initial pairs of states that is required in order to make the measurement independent of the initial state. In the case of a dephasing dynamics the analytical expression of the maximized trace distance is known and it is equal to the absolute value of the dephasing factor [5, 14]

$$C(t) = |\langle e^{-2i\varphi(t)} \rangle|. \quad (4.8)$$

When we consider noisy dephasing dynamics arising from the RTN, the NM behavior of the map changes from Markovian for large values of the switching rate, to non-Markovian for small values of γ . On the contrary, OU-generated dynamics are always Markovian. Being able to simulate the evolution of a quantum system and to control its degree of non-Markovianity or the transition between a Markovian to a NM dynamics, is a very important step in the building and implementation of efficient protocols for quantum communications and quantum information technologies [15–17].

4.3 The Physical Implementation

In order to simulate experimentally this dephasing dynamics we use an all-optical setup [3, 4]. In Fig. 4.1 we show a schematic diagram of the experimental apparatus. The frequency-entangled two-photon state is generated by parametric down-conversion (PDC). We use a diode pump laser @ 405.5 nm, that is temperature stabilized and generates 40 mW @ 70 mA, and a BBO crystal 1 mm thick. The two photons are then collected by two fiber couplers and sent respectively into a single-spatial-mode and polarization-preserving fiber (SMF) and a multimode fiber (MMF). When the idler photon enters the coupler, it travels entirely through the fiber towards the single photon detector (D2). Conversely, the signal photon, after a short fiber (SMF), enters a 4F system [18], i.e. propagates in the air, through few optical devices, the gratings G1 and G2 (1714 lines/mm) and lens L1 and L2 ($f = 500$ mm), an half-wave plate (H1), that we use for the input state preparation, a spatial light modulator (SLM) and a tomographic apparatus (T) to reconstruct the output state. The SLM is a 1D liquid crystal mask (640 pixels, 100 $\mu\text{m}/\text{pixel}$) and is placed on the Fourier plane of the 4f system, between the two lenses L1 and L2 (see Fig. 4.1), where the spectral components of the signal photon are linearly dispersed and focalized (1.82 nm/mm, waist = 60 μm). At the end of the 4F system the signal photon is coupled to a multimode fiber and reaches the single photon detector (D1). The two single photon detectors (D1, D2) are based on the avalanche photodiode C30921S in a passive Geiger mode configuration. Finally an electronic device based on the Ortec 567 TAC/SCA measures the coincidence counts (CC) and sends them to the computer (PC) via the National Instruments card PCI6602. The tomographic apparatus (T) [19, 20] is composed of a quarter-wave plate (Q), an half-wave plate (H) and a polarizer (P). The SLM is controlled by the computer (PC) and is used to introduce a different phase $\varphi_r(t)$ for each pixel.

The experimental setup described above implements our quantum simulator by performing the following steps: the quantum information carrier is a single photon generated by the parametric down-conversion and its polarization is used to encode the state of a qubit; its spectral components are instead exploited to simulate a classical fluctuating field. Specifically, the realizations of the stochastic processes are achieved

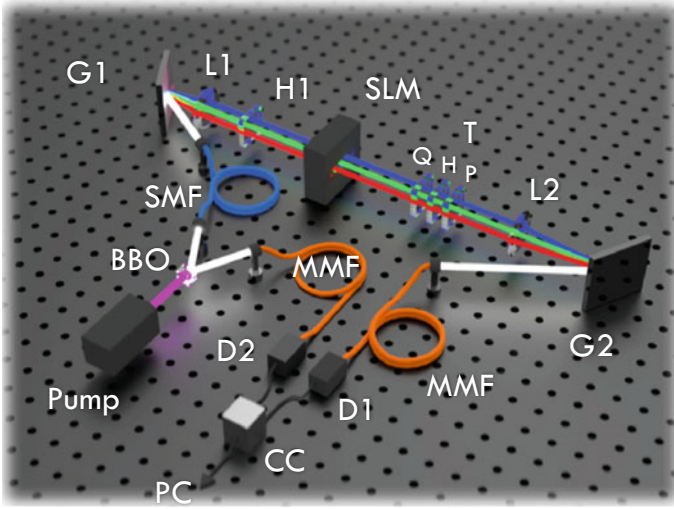


Fig. 4.1 Schematic diagram of our experimental setup. Pump, 405.5 nm laser diode; BBO, Beta barium borate nonlinear crystal; SMF, single-spatial-mode and polarization preserving fiber; MMF, multimode fiber; G1–G2, gratings; L1–L2, lens; H1, half-wave-plate; SLM, spatial light modulator; T, tomographic apparatus; D1–D2, single photon detectors; CC, coincidences counter

in a single shot by using the programmable SLM on the different spectral components of the same photon. The photon is initialized in the state $|\psi_0\rangle = \frac{1}{\sqrt{2}}(|H\rangle + |V\rangle)$ by the H1 plate. By exploiting the quantum nature of the photon, we are able to simulate in one shot the ensemble average over n realizations of the noise. This is accomplished by first numerically simulating on a computer the different n trajectories of the random phases $\varphi(t)$ appearing in (4.3) which are applied only on the horizontal component of the photon $|\psi_r(t)\rangle = \frac{1}{\sqrt{2}}(e^{-2i\varphi_r(t)}|H\rangle + |V\rangle)$. The dynamics $|\psi_r(t)\rangle$ up to time t is obtained using the computer-controlled SLM: each spectral component of the photon passes through a different pixel and acquire a different phase $\varphi(t)$ determined by the computer according to the realizations of the simulated stochastic process. The second step consists in performing the ensemble average by recollecting the different components with a multimode optical fiber, as schematically shown in Fig. 4.1.

We remark here that the qubit system is described by the polarization degree of freedom of the photon, while the spectral degrees of freedom are effectively treated as the environment and will be traced out. Each spectral component is associated to a spatial direction $|x\rangle = |\omega(x)\rangle$ and if we introduce the notation $|\eta_r\rangle$ to describe the r th pixel satisfying the completeness relation, we have $|x\rangle = \sum_r \eta_r(x)|\eta_r\rangle$, where $\eta_r(x) = \langle\eta_r|x\rangle$, that is the projection of the component x on the r th pixel. We assume that the global system is initially in a factorized state and that the polarization is initialized in the state $|D\rangle = \frac{1}{\sqrt{2}}(|H\rangle + |V\rangle)$:

$$\rho_{SE}(0) = \rho_S(0) \otimes \rho_E = |D\rangle\langle D| \otimes \int d\omega |f(\omega)|^2 |\omega\rangle\langle\omega|, \quad (4.9)$$

which can be rewritten in terms of the pixel states as:

$$\rho_{SE}(0) = |D\rangle\langle D| \otimes \int dx |f(x)|^2 |x\rangle\langle x| = |D\rangle\langle D| \otimes \sum_{rs} A_{rs} |\eta_r\rangle\langle\eta_s|, \quad (4.10)$$

where

$$A_{rs} = \int dx |f(x)|^2 \eta_r(x) \eta_s^*(x). \quad (4.11)$$

Since on the horizontal component is imposed a pixel-dependent phase $\varphi_r(t)$, we can write the effective global evolution operator as:

$$U(t) = \exp \left[-2i |H\rangle\langle H| \otimes \sum_r \varphi_r(t) |\eta_r\rangle\langle\eta_r| \right], \quad (4.12)$$

which can be rewritten as:

$$U(t) = |H\rangle\langle H| \otimes \sum_r e^{-2i\varphi_r(t)} |\eta_r\rangle\langle\eta_r| + |V\rangle\langle V| \otimes \sum_r |\eta_r\rangle\langle\eta_r|. \quad (4.13)$$

The global evolved state $\rho_{SE}(t) = U(t)\rho_{SE}(0)U^\dagger(t)$ can be written in a compact notation as:

$$\rho_{SE}(t) = \frac{1}{1} \sum_{r,s} A_{rs} \begin{pmatrix} e^{-2i(\varphi_r(t)-\varphi_s(t))} |\eta_r\rangle\langle\eta_s| & e^{-2i\varphi_r(t)} |\eta_r\rangle\langle\eta_s| \\ e^{2i\varphi_r(t)} |\eta_r\rangle\langle\eta_s| & |\eta_r\rangle\langle\eta_s| \end{pmatrix} \quad (4.14)$$

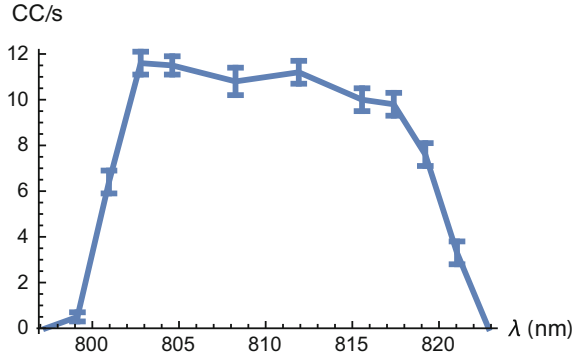
and therefore the marginals are:

$$\rho_S(t) = \frac{1}{2} \sum_r A_{rr} \begin{pmatrix} 1 & e^{-2i\varphi_r(t)} \\ e^{2i\varphi_r(t)} & 1 \end{pmatrix} \quad (4.15)$$

$$\rho_E(t) = \sum_{r,s} A_{rs} \frac{1}{2} (e^{-2i(\varphi_r(t)-\varphi_s(t))} + 1) |\eta_r\rangle\langle\eta_s|. \quad (4.16)$$

Both the system and the environment are subject to a dephasing dynamics, where the populations do not evolve, while the off-diagonal elements are function of time. Let us focus on the expression of the system density matrix $\rho_S(t)$. Since the PDC spectrum $f(x)$ is almost constant in the considered range of values (cf. Fig. 4.2), the coefficients $A_{rr} = \int dx |f(x)|^2 |\eta_r(x)|^2 = n^{-1}$ become constant. Here n is the total number of pixels considered. It follows that the system density matrix becomes the one in (4.3), with

Fig. 4.2 The measured spectrum of the PDC. We can see that it is almost flat in the region 802–817 nm



$$\langle e^{2i\varphi(t)} \rangle_n = \frac{1}{n} \sum_{r=1}^n e^{2i\varphi_r(t)}. \quad (4.17)$$

This description is ideal and it does not consider the effects of imperfection and noise in the experimental apparatus. In order to give a more realistic representation of the system dynamics, we take into account the fact that the initial state of the system is not exactly pure, but a mixed state with high purity $\rho_{S,\text{exp}}(t) = p\rho_S(t) + (1-p)\rho_{\text{mix}}$, where the parameter p is related to the purity of the initial state, with $p = 1$ corresponding to the pure state $\rho_S = |D\rangle\langle D|$ and $p = 0$ to the maximally mixed state $\rho_{\text{mix}} = \frac{1}{2}(|H\rangle\langle H| + |V\rangle\langle V|)$. The system density matrix (4.15) can now be rewritten as:

$$\rho_{S,\text{exp}}(t) = \frac{1}{2} \begin{pmatrix} 1 & p \langle e^{-2i\varphi(t)} \rangle_n \\ p \langle e^{2i\varphi(t)} \rangle_n & 1 \end{pmatrix} \quad (4.18)$$

where the off-diagonal elements are now multiplied by the factor p . Equation (4.18) describes a dephasing map for a qubit. The off-diagonal elements of such evolutions, in addition to representing the coherences of the quantum state, are connected to the non-Markovianity of the dynamical map [5], which is nowadays considered a resource for quantum technologies [21]. For this reason we are interested in reconstructing the off-diagonal element of the density matrix and we assume that this value is a real quantity. Specifically, we want to reconstruct the maximized trace distance in (4.8). This can be experimentally done in two ways. The first approach consists in reconstructing the whole density matrix through qubit state tomography [19]. This means performing four projective measurements and to use their outputs to build a maximum likelihood estimator for the elements of the density matrix. The second approach exploits the fact that the off-diagonal coefficient is real to make only one projective measurement on the system. In particular, if we project onto the state $|D\rangle$, we obtain:

$$\langle D | \rho_{S,\text{exp}} | D \rangle = \frac{1}{2} (1 + p \Re \langle e^{-2i\varphi(t)} \rangle_n) \quad (4.19)$$

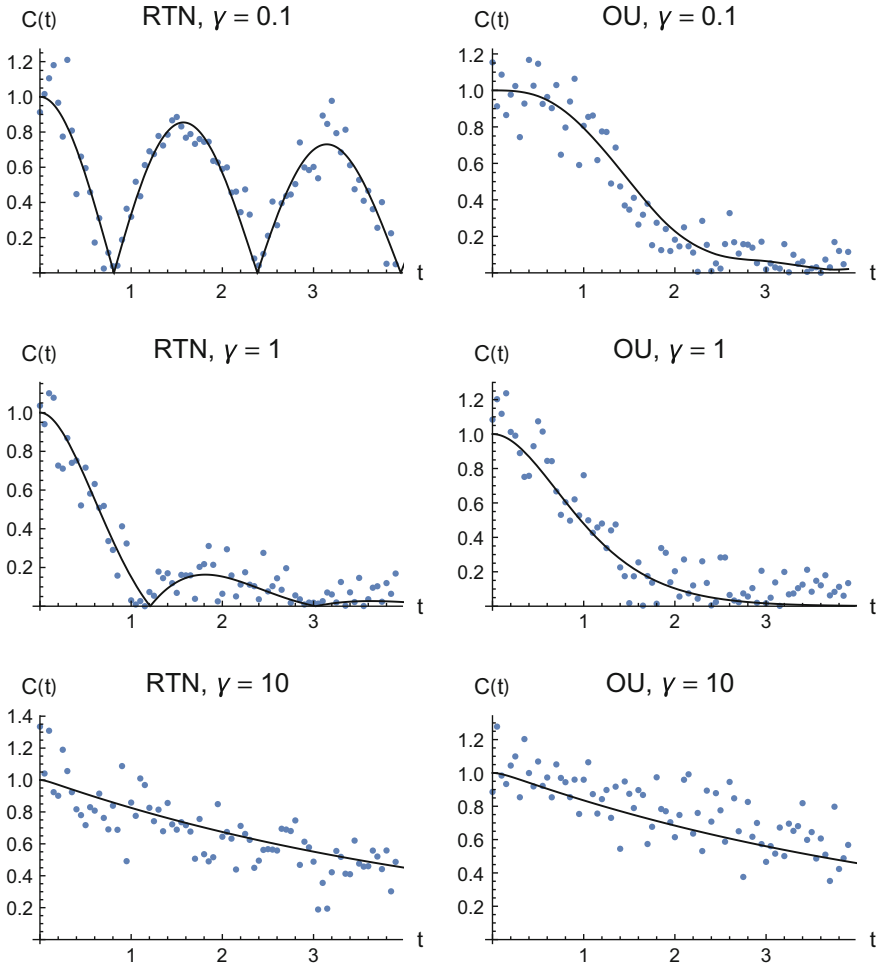


Fig. 4.3 Dynamics of the absolute value of the off-diagonal element of $\rho_S(t)$, $C(t) = |\langle e^{-2i\varphi_r(t)} \rangle_n|$, for RTN (left) and OU (right), for $\gamma = 0.1$, $\gamma = 1$ and $\gamma = 10$ (from top to bottom). The blue dots are obtained experimentally by projection onto the state $|+\rangle$, corrected with the factor p , as discussed in the main text. The black line is the analytic solution of the model

which can easily give us the wanted quantity $C(t) = |\langle e^{-2i\varphi(t)} \rangle|$ by means of a single measurement. In order to obtain the experimental value of p we performed a reference measure in the presence of static noise (we considered the RTN with $\gamma = 0$) as described in [3].

It is worth noting that, because of the limited number of pixels of our SLM, the ensemble average is realized over a fixed number $n = 100$ of realizations of the fluctuating random fields. However, we show that this number is sufficient to correctly reproduce the theoretical behavior expected for NM, as can be seen in

Fig. 4.3, where we compare the simulated dynamics with the theoretical predictions. By changing the value of the switching rate γ , we are able to simulate the dynamics of the maximized trace distance $C(t)$ for the two different processes here considered. According to the theoretical model, we can move from a non-Markovian dynamics to a Markovian one for the RTN case by increasing the value of γ . Indeed, small values of γ lead to oscillations in the maximized trace distance, and the amplitude of these revivals is larger for smaller values of the switching rate. When the γ is above a certain threshold [5], the oscillations fade away and decoherence factor decays monotonically in time, meaning that memory effects do not play a role anymore in the dynamics. The OU process, instead, always leads to a monotonic behavior for $C(t)$, a signature of a Markovian quantum dynamics where information cannot flow back to the system.

4.4 Conclusions

We have suggested and demonstrated a quantum simulator based only on optical elements. This simulator can mimic the dynamics of single qubits in dephasing channels stemming from the interaction with classical fluctuating fields. The qubit is encoded in the polarization degrees of freedom of a single photon generated by PDC, while the several realizations of the classical noise are instead implemented on its spectral components by means of a programmable SLM. As proof of principle, we have shown that our simulator allows us to simulate quantum dynamics with different degrees of non-Markovianity, evaluated in terms of backflow of information from the environment to the system.

It is worth noting that our simulator can be also used to analyze the relation between the classical non-Markovianity of a classical fluctuating field and the non-Markovianity of the induced quantum dynamical map, by taking into account other kinds of stochastic processes. This will help to shed light into the open problem of finding a connection between the two definitions of non-Markovianity, used in the classical and quantum scenarios.

References









1. R.P. Feynman, Simulating physics with computers. *Int. J. Theor. Phys.* **21**, 467 (1982)
2. T.H. Johnson, S.R. Clark, D. Jaksch, What is a quantum simulator? *EPJ Quant. Technol.* **1**, 10 (2014)
3. S. Cialdi, M.A.C. Rossi, C. Benedetti, B. Vacchini, D. Tamascelli, S. Olivares, M.G.A. Paris, All-optical quantum simulator of qubit noisy channels. *Appl. Phys. Lett.* **110**, 081107 (2017)
4. M.A.C. Rossi, C. Benedetti, S. Cialdi, D. Tamascelli, S. Olivares, B. Vacchini, M.G.A. Paris, Non-Markovianity by undersampling in quantum optical simulators. *Int. J. Quantum Inf.* **15**, 1740009 (2017)
5. C. Benedetti, M.G.A. Paris, S. Maniscalco, Non-Markovianity of colored noisy channels. *Phys. Rev. A* **89**, 012114 (2014)

6. M.A.C. Rossi, C. Benedetti, M.G.A. Paris, Engineering decoherence for two-qubit systems interacting with a classical environment. *Int. J. Quant. Inf.* **12**, 1560003 (2014)
7. C. Benedetti, M.G.A. Paris, Effective dephasing for a qubit interacting with a transverse classical field. *Int. J. Quant. Inf.* **12**, 1461004 (2014)
8. M.A.C. Rossi, M.G.A. Paris, Non-Markovian dynamics of single- and two-qubit systems interacting with Gaussian and non-Gaussian fluctuating transverse environments. *J. Chem. Phys.* **144**, 024113 (2016)
9. H.J. Wold, H. Brox, Y.M. Galperin, J. Bergli, Decoherence of a qubit due to either a quantum fluctuator, or classical telegraph noise. *Phys. Rev. B* **86**, 205404 (2012)
10. A. Rivas, S.F. Huelga, M.B. Plenio, Entanglement and Non-Markovianity of quantum evolutions. *Phys. Rev. Lett.* **105**, 050403 (2010)
11. S. Lorenzo, F. Plastina, M. Paternostro, Geometrical characterization of non-Markovianity. *Phys. Rev. A* **88**, 020102(R) (2013)
12. X.-M. Lu, X. Wang, C.P. Sun, Quantum Fisher information flow and non-Markovian processes of open systems. *Phys. Rev. A* **82**, 042103 (2010)
13. H.-P. Breuer, E.M. Laine, J. Piilo, Measure for the degree of non-Markovian behavior of quantum processes in open systems. *Phys. Rev. Lett.* **103**, 210401 (2009)
14. Z. He, J. Zou, L. Li, B. Shao, Effective method of calculating the non-Markovianity N for single-channel open systems. *Phys. Rev. A* **83**, 012108 (2011)
15. A.W. Chin, S.F. Huelga, M.B. Plenio, Quantum metrology in non-Markovian environments. *Phys. Rev. Lett.* **109**, 233601 (2012)
16. R. Vasile, S. Olivares, M.G.A. Paris, S. Maniscalco, Continuous variable quantum key distribution in non-Markovian channels. *Phys. Rev. A* **83**, 042321 (2011)
17. S.F. Huelga, A. Rivas, M.B. Plenio, Non-Markovianity-assisted steady state entanglement. *Phys. Rev. Lett.* **108**, 160402 (2012)
18. A.M. Weiner, Femtosecond pulse shaping using spatial light modulators. *Rev. of Sci. Instrum.* **71**, 1929 (2000)
19. D.F.V. James, P.G. Kwiat, W.J. Munro, A.G. White, Measurement of qubits. *Phys. Rev. A* **64**, 052312 (2001)
20. K. Banaszek, G.M. D'Ariano, M.G.A. Paris, M.F. Sacchi, Maximum-likelihood estimation of the density matrix. *Phys. Rev. A* **61**, 010304(R) (1999)
21. B. Bylicka, D. Chruściński, S. Maniscalco, Non-Markovianity and reservoir memory of quantum channels: a quantum information theory perspective. *Sci. Rep.* **4**, 5720 (2014)

Chapter 5

Innovative Instrumentation for the Study of Atmospheric Aerosol Optical Properties



Vera Bernardoni , Alice C. Forello , Federico Mariani, Bruno Paroli, Marco A. C. Potenza , Alberto Pullia , Francesco Riccobono, Tiziano Sanvito , Sara Valentini , Gianluigi Valli  and Roberta Vecchi 

Abstract Aerosol optical properties (i.e. scattering and absorption) are of great importance to assess aerosol effects e.g. on visibility and Earth radiation balance. In this paper, we present innovative optical instrumentation developed at the Department of Physics “Aldo Pontremoli” of the University of Milan: a multi-wavelength polar photometer (PP_UniMI) and a Single Particle Extinction and Scattering (SPES). PP_UniMI is a filter-based device providing the aerosol absorption coefficient of aerosol at 4 wavelengths (λ). Such measurements are of interest to have insights into the λ -dependent behavior of aerosol absorption properties, which is still poorly

V. Bernardoni (✉) · A. C. Forello · F. Mariani · B. Paroli · M. A. C. Potenza · A. Pullia · T. Sanvito
S. Valentini · G. Valli · R. Vecchi
Department of Physics “Aldo Pontremoli”, Università degli Studi di Milano, Via Celoria 16,
20133 Milan, Italy
e-mail: vera.bernardoni@unimi.it

A. C. Forello
e-mail: alicecorina.forello@studenti.unimi.it

F. Mariani
e-mail: federico.mariani@eosinstruments.com

B. Paroli
e-mail: bruno.paroli@unimi.it

M. A. C. Potenza
e-mail: marco.potenza@unimi.it

A. Pullia
e-mail: alberto.pullia@unimi.it

T. Sanvito
e-mail: tiziano.sanvito@eosinstruments.com

© Springer Nature Switzerland AG 2018
P. F. Bortignon et al. (eds.), *Toward a Science Campus in Milan*,
https://doi.org/10.1007/978-3-030-01629-6_5

understood especially for what concerns weakly absorbing aerosol components. SPES allows to determine the size and refractive index of single particles. In case of absorbing particles, also information on the imaginary part of the refractive index—very important especially in the field of global models—can be provided with little assumptions. We describe the main features of the two instruments and their advantages and limitations. Examples of application are also presented.

5.1 Introduction

Light scattering and absorption properties by atmospheric aerosol are responsible for the aerosol effects e.g. on visibility impairment at local scale or the Earth radiation balance at global scale. The main physical parameters allowing the description of light scattering and absorption by a particle are its size, complex refractive index, mixing state, and shape [1].

Great uncertainty is still associated to the estimate of the aerosol-radiation effect on the Earth radiation balance [2] and refining the parameters used as input in climate models is mandatory for a reduction of such uncertainties. Aerosol absorption properties are of peculiar interest as they can provide positive contribution to the Earth radiation balance, but many aspects are not fully understood yet. Indeed, different species can absorb visible light. The main light absorbing component in atmospheric aerosol is recognized to be the black carbon (BC). Besides BC, in the last fifteen years, it was evidenced that there exist organic compounds weakly absorbing at long visible wavelengths (λ), but with enhanced absorption behavior at low visible and near-UV wavelengths (called brown carbon—BrC) [3]. BrC chemical properties, sources, and light absorption characteristics are still poorly investigated [4, and therein cited literature]. Consequently, it is currently barely accounted for in global models. Thus, developing instrumentation to gain insight into the parameters

S. Valentini

e-mail: sara.valentini@unimi.it

G. Valli

e-mail: gianluigi.valli@unimi.it

R. Vecchi

e-mail: roberta.vecchi@unimi.it

V. Bernardoni · A. C. Forello · B. Paroli · M. A. C. Potenza · A. Pullia · S. Valentini · G. Valli

R. Vecchi

INFN-Milan, Via Celoria 16, 20133 Milan, Italy

F. Riccobono

Joint Research Centre, European Commission, Via Fermi 2749, 21027 Ispra, Italy

e-mail: fraricco@gmail.com

T. Sanvito

EOS s.r.l., Viale Ortles 22/4, 20139 Milan, Italy

influencing the aerosol absorption properties and/or into the λ -dependent absorption coefficient in atmosphere is mandatory to better understand the role of atmospheric aerosol on the Earth radiation balance. As an example, it was evidenced that small variation on the imaginary part of the refractive index can have important impact on the modeled radiative forcing [5].

As the particle mixing state plays a key role in the aerosol-radiation interaction processes, it is important to perform measurement of light scattering and absorption properties of atmospheric particles with no sample pre-treatment. To this aim, two instruments were developed for the study of aerosol optical properties at the Department of Physics “Aldo Pontremoli” of the University of Milan: the multi- λ polar photometer PP_UniMI [6, 7] and the Single Particle Extinction and Scattering instrument [8].

PP_UniMI was originally developed to measure off-line the aerosol absorption coefficient at 635 nm on aerosol collected on 47-mm filters (i.e. filters commonly used during sampling campaigns) with no need of additional instrumentation. In this configuration, the agreement with an external reference instrument—the Multi-Angle Absorption Photometer (MAAP)—was within 2% [6]. Recently, implementations were realized to perform measurements at 4 λ s and to extend the application of the methodology to samples collected with 1-h resolution [7]. In both cases, no sample pre-treatment is required. Multi- λ measurements of aerosol absorption properties can be exploited to gain insight into BrC properties applying optical source and component apportionment models (e.g. Multi-Wavelength Absorption Analyzer model—MWA [9, 10]).

The SPES concept was recently introduced and validated for liquid suspensions [11] and in the following extended to particles suspended in air [8]. SPES is a single-particle device providing information on the real and imaginary parts of the adimensional scattering amplitude [12] according to the optical theorem [1, 13]. As a consequence, particle characteristics (i.e. particle size and refractive index) can be provided with few or no assumptions [14, 15].

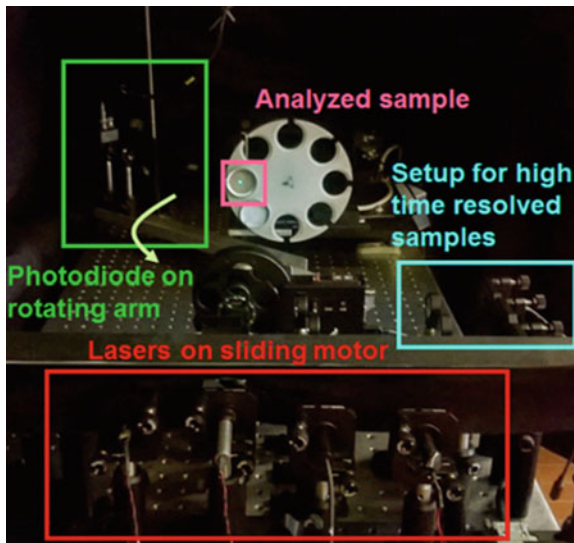
The features of PP_UniMI and SPES were presented at the Congress of the Physics of the University of Milan in June 2017 together with some applications. In this paper, the main contents of the presentation will be reported.

5.2 Description of the Instruments

5.2.1 The Polar Photometer PP_UniMI

PP_UniMI provides information on the aerosol absorption coefficient at four wavelengths (405, 532, 635, 780 nm) starting from measurements of the intensity of the light transmitted and scattered in the front and back hemispheres by a blank and aerosol-loaded filter. The chosen laser beam impinges on the sample, and the scan of the scattering plane (0–173° with about 0.4° resolution) is performed thanks to a

Fig. 5.1 PP_UniMI set-up



photodiode mounted on a rotating arm (see Fig. 5.1). Solid angle integration allows to determine the total amount of light diffused in the two hemispheres, and these data are the input to the radiative transfer model developed by Hänel [16, 17] for membrane filters and adapted by Petzold and Schönlinner [18] for the application to aerosol collected on fiber filters. The outputs of the model are the single scattering albedo (ω) and the optical depth (τ) of the layer containing the particles. Considering the sampled volume (V) and the deposit area (A), the aerosol absorption coefficient (b_{abs} , expressed in M m^{-1}) in atmosphere during the sampling is determined as:

$$b_{\text{abs}} = (1 - \omega)\tau \frac{A}{V}$$

PP_UniMI was recently upgraded to analyze samples collected with high temporal resolution using a streaker sampler [19], characterized by low loadings and collected on very thin substrates. Multiple scattering between the sample and an additional scattering matrix leaned against the sample were exploited to increase the instrument sensitivity. Good agreement with an external reference instrument—the Multi-Angle Absorption Photometer MAAP [18]—was evidenced (slope = 1.10, $R^2 = 0.93$). Further details can be found in Bernardoni et al. [7].

5.2.2 Single Particle Extinction and Scattering (SPES)

SPES allows gaining information on the real and imaginary parts of the forward complex scattering amplitude thanks to a self-interferometric scheme by exploiting

the optical theorem (OT) [1, 13]. Currently, single wavelength measurements are performed at 640 nm.

Briefly, a particle crosses a beam waist perpendicularly. When the particle is exactly in the center of the beam, an attenuation figure is monitored onto a segmented quadrant photodiode (QPD) placed in the far field. The OT is used to relate the extinction cross section to the real part of the forward scattering amplitude. When the particle is slightly displaced from the optical axis of the system, transmitted and scattered radiation combine forming an interference figure: a small, asymmetric signal fluctuation is generated on the QPD [11], and it is related to the imaginary part of the forward scattering amplitude. Thus, combining the two measurements gives access to the whole complex forward scattering amplitude.

Acquisition of the QPD signals is performed by a custom, front-end electronics properly developed. It separates the signal into two components: the slow signal is isolated using a low-pass filter, allowing monitoring the laser intensity continuously; then a fast, zero-average signal corresponding to the rapid fluctuations related to the particle crossing the beam waist is digitized at 12-bit, thus allowing detailed signal analysis [20].

Among the advantages of the technique, it is noteworthy that:

- No calibration is required: the Mie theory is exploited to relate the complex forward scattering amplitude to the particle characteristics;
- Spurious signals are automatically rejected thanks to pulse-shape analysis;
- Information on the imaginary part of the refractive index can be obtained under few assumptions or performing size-segregation upstream SPES;
- Information on non-spherical particles can be obtained.

Further details on the SPES set-up for aerosol suspended in air can be found in Mariani et al. [8].

5.3 Results and Discussion

5.3.1 Example of Results by PP_UniMI

As an example of PP_UniMI application, 4- λ analysis of the absorption coefficient of aerosol collected using a streaker sampler is shown in Fig. 5.2. Aerosol was collected using a streaker sampler at an urban background site in Milan [see e.g. 7] in December 2016.

It is noteworthy that multi- λ analysis can be exploited to perform studies of optical source apportionment, using literature approaches—e.g. the Multi-Wavelength Absorption Analyzer model, which provides both source and component apportionment of the aerosol absorption coefficient at different λ s [9].

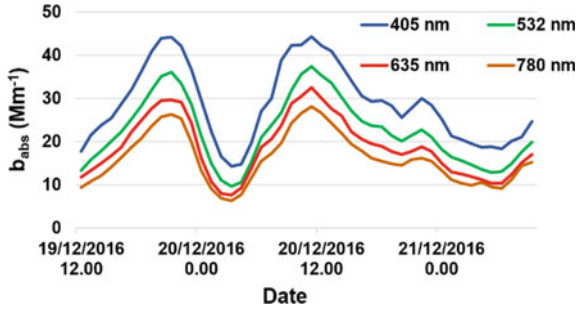


Fig. 5.2 Example of 4- λ measurements of the aerosol absorption coefficient

5.3.2 Examples of Results by SPES

SPES was tested on laboratory aerosol: water droplets, NaCl particles, and graphite nanoparticles. In Fig. 5.3 SPES measurements for each aerosol type are represented as two dimensional histograms showing the number of events recorded in each two dimensional bin in the $S(0)$ complex plane. Populations are characterized by size polydispersity, giving the main elongation of the data sets. Data exhibit different spreads along the imaginary axis, the smaller indicating higher homogeneity and sphericity. Changes in the complex amplitudes of the fields scattered by different aerosol types are evident.

SPES showed very good performances on test aerosol: for pure water droplets generated using an aerosol generator—ATM 220 by Topas—the retrieved real refractive index was 1.32 ± 0.01 , in close agreement with the expected value (1.33). The maximum detected particle diameter was $d \sim 1 \mu\text{m}$, consistently with the atomizer cut-off diameter. The small spread of the measurements in the scattering plane (Fig. 5.3a) is indication of particle homogeneity and sphericity.

The same experimental set-up was used to generate and characterize NaCl particles. Spread of the data in the complex scattering plane (Fig. 5.3b) and consequent broader distribution of the retrieved real refractive index gave indication of non-spherical particles. The retrieved NaCl size distribution had a sudden drop at $d < 600 \text{ nm}$. It is noteworthy that SPES provides geometric diameters (i.e. diameters of a spherical particle with the same surface area as the measured one), whereas the aerosol generator cut-off refers to aerodynamic diameter (i.e. diameter of a spherical particle with density $\rho = 1 \text{ g/cm}^3$ having the same inertial properties of the measured particle). Considering the NaCl density (2.16 g/cm^3), the expected geometric cut-off diameter of the aerosol generator assuming spherical particles is 680 nm, which is consistent with the sudden drop observed in NaCl particle size distribution.

Graphite nanoparticles (proxy for black carbon) were also analyzed by SPES: they were produced by a graphite spark aerosol generator by Palas GmbH (Mod. DNP digital 3000). Graphite nanoparticles are expected to have high imaginary part of the refractive index. Furthermore, they are expected as agglomerates (i.e. far from

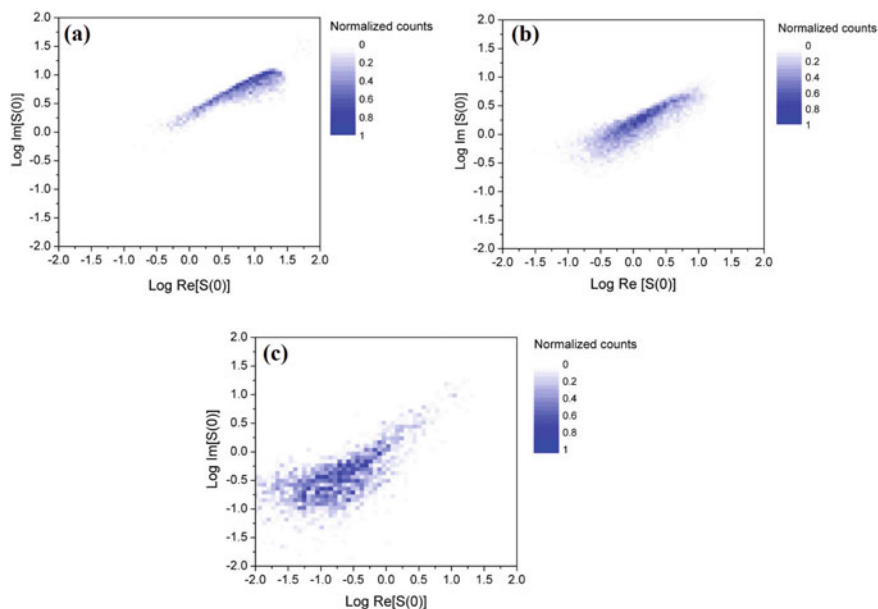


Fig. 5.3 SPES results represented as two dimensional histograms showing the number of events recorded in each two dimensional bin in the $S(0)$ complex plane. **a** Water droplets; **b** NaCl particles; **c** graphite nanoparticles

spherical shape). SPES measurement distribution is spread throughout the complex plane, as expected for non-spherical, inhomogeneous particles. Furthermore, measurements are generally placed in an area with lower imaginary part of the forward scattering function, which is an indication of absorptive particles (Fig. 5.3c). Calculation by the Mie theory was performed—even if inaccuracies in the results are expected due to the poor consistency of the graphite particles characteristics with the hypotheses of the theory i.e. particle sphericity and homogeneity—and the effective imaginary part of the refractive index resulted to be $k = 0.2$.

5.3.3 *PP_UniMI and SPES Joint Application for the Characterization of Unknown Material*

PP_UniMI and SPES can be separately used for aerosol characterization. Nevertheless, they can be used jointly for a more detailed characterization of unknown materials. This was performed by driving aerosol through SPES and sampling the aerosol at the instrument outlet on PTFE filters. The aerosol collected on filter was then analyzed by PP_UniMI. This procedure was performed to characterize a test material of unknown physical-chemical and optical features (pyrethrum smoke).

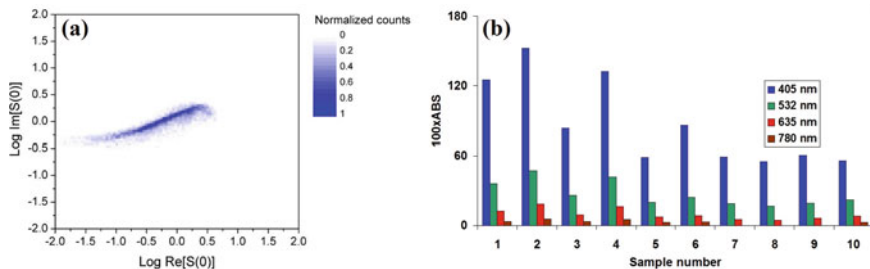


Fig. 5.4 a SPES results represented as in Fig. 5.3. b $4\text{-}\lambda$ sample absorbance measured by PP_UniMI

SPES results showed a very small spread of the data along the imaginary axis in the complex plane (Fig. 5.4a), thus providing indication for homogeneous, spherical particles. The evident change of the population shape occurring in the left-hand side of the plot indicates a modification of the overall structure of the particles. Pyrethrum smoke particles resulted to be weakly absorbing at 640 nm (imaginary part of the refractive index $k = 0.12$) and their size was in the range 200–700 nm diameter.

Furthermore, PP_UniMI measurements provided information on the λ -dependence of pyrethrum smoke sample absorbance (Fig. 5.4b). In general, it is expected a $\lambda^{-\alpha}$ dependence for the aerosol absorption coefficient, where α is called Ångström absorption exponent. α is expected to be 1 for pure BC particles in the Rayleigh regime, whereas it is expected to be much higher for BrC [e.g. 4, 21, 22, and therein cited literature]. Pyrethrum smoke showed $\alpha = 5$: as carbonaceous particles are expected in Pyrethrum smoke, our measurements suggest that they consist mainly of BrC (further details in Mariani et al. [8]).

5.4 Conclusions and Perspectives

Optical instruments for the characterization of atmospheric aerosol (PP_UniMI, SPES) were developed at the Department of Physics “Aldo Pontremoli” of the University of Milan. These instruments provide information useful to fill the gap of knowledge currently present in the aerosol absorption and scattering properties and in the evaluation of their effect, especially at global scale. Both instruments were developed to analyze samples with no need of pre-treatment, which is important to avoid modification of the optical properties related to particle mixing state.

The polar photometer PP_UniMI provides multi- λ measurements of the absorption coefficient of aerosol collected on filters. This information can be exploited in source and component apportionment studies [9] to gain information on BrC properties. Reliability of the methodology was successfully tested against independent measurements performed by a Multi-Angle Absorption Photometer (MAAP) at 635 nm.

The Single Particle Extinction and Scattering instrument (SPES) provides information on the size and real refractive index of non-absorbing particles. Furthermore, information on the imaginary part of the particle refractive index can be obtained with few assumptions. Single particle measurements are of great importance as they can both allow to study particle ensemble characteristics and to focus on possible subsets of interest.

Testing the joint application of the two instruments to measure particles of unknown material showed potentiality of providing very important information for the characterization of the size, shape, and λ -dependent light absorption behavior of the tested particles. This opens in perspective the possibility of better characterizing other test materials, and possibly ambient aerosol samples.

Acknowledgements The authors are grateful to F. Cavaliere and D. Viganò of the Mechanical Workshop of the Physics Department for technical support.

References

1. C.F. Bohren, D.R. Huffman, *Absorption and Scattering by Small Particles* (Wiley, New York, 1983)
2. T.F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, P.M. Midgley (eds.), *IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013), 1535 pp
3. M.O. Andreae, A. Gelencsér, Black carbon or brown carbon? The nature of light-absorbing carbonaceous aerosols. *Atmos. Chem. Phys.* **6**, 3131–3148 (2006). <https://doi.org/10.5194/acp-6-3131-2006>
4. A. Laskin, J. Laskin, S.A. Nizkorodov, Chemistry of atmospheric brown carbon. *Chem. Rev.* **115**, 4335–4382 (2015). <https://doi.org/10.1021/cr5006167>
5. K.J. Zarzana, C.D. Cappa, M.A. Tolbert, Sensitivity of aerosol refractive index retrievals using optical spectroscopy. *Aerosol Sci. Technol.* **48**, 1133–1144 (2014). <https://doi.org/10.1080/02786826.2014.963498>
6. R. Vecchi, V. Bernardoni, C. Paganelli, G. Valli, A filter-based light-absorption measurement with polar photometer: effects of sampling artefacts from organic carbon. *J. Aerosol Sci.* **70**, 15–25 (2014). <https://doi.org/10.1016/j.jaerosci.2013.12.012>
7. V. Bernardoni, G. Valli, R. Vecchi, Set-up of a multi wavelength polar photometer for off-line absorption coefficient measurements on 1-h resolved aerosol samples. *J. Aerosol Sci.* **107**, 84–93 (2017). <https://doi.org/10.1016/j.jaerosci.2017.02.009>
8. F. Mariani, V. Bernardoni, F. Riccobono, R. Vecchi, G. Valli, T. Sanvito, B. Paroli, A. Pullia, M.A.C. Potenza, Single particle extinction and scattering allows novel optical characterization of aerosols. *J. Nanopart. Res.* **19**, 291 (2017). <https://doi.org/10.1007/s11051-017-3995-3>
9. D. Massabò, L. Caponi, V. Bernardoni, M.C. Bove, P. Brotto, G. Calzolari, F. Cassola, M. Chiari, M.E. Fedi, P. Fermo, M. Giannoni, F. Lucarelli, S. Nava, A. Piazzalunga, G. Valli, R. Vecchi, P. Prati, Multi-wavelength optical determination of black and brown carbon in atmospheric aerosols. *Atmos. Environ.* **108**, 1–12 (2015). <https://doi.org/10.1016/j.atmosenv.2015.02.058>
10. V. Bernardoni, R.E. Pileci, L. Caponi, D. Massabò, The multi-wavelength absorption analyzer (MWAA) model as a tool for source and component apportionment based on aerosol absorption properties: application to samples collected in different environments. *Atmosphere* **8**, 218 (2017). <https://doi.org/10.3390/atmos8110218>

11. M.A.C. Potenza, T. Sanvito, A. Pullia, Measuring the complex field scattered by single submicron particles. *AIP Adv.* **5**, 117222 (2015). <https://doi.org/10.1063/1.4935927>
12. M.A.C. Potenza, P. Milani, Free nanoparticle characterization by optical scattering field analysis: opportunities and perspectives. *J. Nanopart. Res.* **16**, 2680 (2014). <https://doi.org/10.1007/s11051-014-2680-z>
13. H.C. Van de Hulst, *Light Scattering by Small Particles* (Dover Publication Inc., New York, 1957)
14. M.A.C. Potenza, T. Sanvito, S. Argentiere, C. Cella, B. Paroli, C. Lenardi, P. Milani, Single particle optical extinction and scattering allows real time quantitative characterization of drug payload and degradation of polymeric nanoparticles. *Sci. Rep.* **5**, 18228 (2015). <https://doi.org/10.1038/srep18228>
15. S. Villa, T. Sanvito, B. Paroli, A. Pullia, B. Delmonte, M.A.C. Potenza, Measuring shape and size of micrometric particles from the analysis of the forward scattered field. *J. Appl. Phys.* **119**, 224901 (2016). <https://doi.org/10.1063/1.4953332>
16. G. Hänel, Radiation budget of the boundary layer: Part II. Simultaneous measurement of mean solar volume absorption and extinction coefficients of particles. *Beiträge zur Physik der Atmosphäre* **60**, 241–247 (1987)
17. G. Hänel, Optical properties of atmospheric particles: complete parameter sets obtained through polar photometry and an improved inversion technique. *Appl. Opt.* **33**, 7187–7199 (1994). <https://doi.org/10.1364/AO.33.007187>
18. A. Petzold, M. Schönlinner, Multi-angle absorption photometry—a new method for the measurement of aerosol light absorption and atmospheric black carbon. *J. Aerosol Sci.* **35**, 421–441 (2004). <https://doi.org/10.1016/j.jaerosci.2003.09.005>
19. H.J. Annegarn, Time-series analysis of PIXE aerosol measurements. *Nucl. Instrum. Methods Phys. Res. B* **22**, 270–274 (1987). [https://doi.org/10.1016/0168-583X\(87\)90341-7](https://doi.org/10.1016/0168-583X(87)90341-7)
20. A. Pullia, T. Sanvito, M.A. Potenza, F. Zocca, A low-noise large dynamic-range readout suitable for laser spectroscopy with photodiodes. *AIP Rev. Sci. Instrum.* **83**, 104704 (2012). <https://doi.org/10.1063/1.4756045>
21. N. Utry, T. Ajtai, Á. Filep, M. Dániel Pintér, A. Hoffer, Z. Bozoki, G. Szabó, Mass specific optical absorption coefficient of HULIS aerosol measured by a four-wavelength photoacoustic spectrometer at NIR, VIS and UV wavelengths. *Atmos. Environ.* **69**, 321–324 (2013). <https://doi.org/10.1016/j.atmosenv.2013.01.003>
22. D.A. Lack, J.M. Langridge, On the attribution of black and brown carbon light absorption using the Ångström exponent. *Atmos. Chem. Phys.* **13**, 10535–10543 (2013). <https://doi.org/10.5194/acp-13-10535-2013>

Chapter 6

Giotto Unveiled: New Developments in Imaging and Elemental Analysis Techniques for Cultural Heritage



Nicola Ludwig , Letizia Bonizzoni , Michele Caccia, Francesco Cavaliere, Marco Gargano , Daniele Viganò, Christian Salvatore, Matteo Interlenghi, Marco Martini and Anna Galli

Abstract The Giotto's masterpiece God the Father with Angels, never investigated till now, was studied by our team of local researchers, involved in application of scientific methods for cultural heritage since many years. Exploiting the integration of different knowledges, technologies and resources of our team, we were able to provide data to understand the painting technique, the pigment used and the underdrawing of this Giotto's painting. We performed the following non-invasive analyses: Macro-XRF scanning (MA-XRF), Fiber optic reflectance spectroscopy (FORS), high resolution IR scanning reflectography, infrared false color (IRFC). Only portable instrumentations were used, with operating times compatible with the opening hours of exhibition. In particular, the analytical campaign was the opportunity to test the portable IR scanning prototype based on a peculiar spherical scanning system characterized by light weight and low cost motorized head. The analytical results revealed a painting technique already used by Giotto and based on different superimposed pigment layers. By combining the effectiveness of scanning portable-XRF (pXRF) with the responsive of image spectroscopic analysis, we move step by step toward the discovery of Giotto's palette, with particular attention to the flesh tones in God the Father with Angels. The imaging data support the hypothesis of a detailed underlying

N. Ludwig (✉) · L. Bonizzoni · F. Cavaliere · M. Gargano · D. Viganò
Dipartimento di Fisica "Aldo Pontremoli", Università degli Studi di Milano, Milan, Italy
e-mail: nicola.ludwig@unimi.it

M. Caccia · A. Galli
CNR-IFN, Milan, Italy

M. Martini · A. Galli
Dipartimento di Scienza dei Materiali, Università di Milano-Bicocca, Milan, Italy

M. Caccia · C. Salvatore · M. Interlenghi
CNR-IBFM, Segrate, Milan, Italy

© Springer Nature Switzerland AG 2018
P. F. Bortignon et al. (eds.), *Toward a Science Campus in Milan*,
https://doi.org/10.1007/978-3-030-01629-6_6

sketch that includes also a drawing characterized by larger brush signs; the use of patrones for the face of “God” was supposed thanks to comparison with other Giotto masterpieces.

6.1 Introduction

God the Father with Angels (about 1330, tempera on gold background, 76 × 71 cm) is the cusp of the altarpiece polyptych of Baroncelli Chapel painted by Giotto (1267–1337) in the Franciscan church of Santa Croce in Florence (Fig. 6.1). The upper part, representing six angels floating in a heavenly state and paying tribute to God, was later separated from the main altarpiece that is still conserved in Florence. The cusp-shaped masterpiece was finally purchased on a private market of works of art and then donated to the San Diego Museum of Art (CA), where it is presently conserved. Little is known about this painting history since the ill-judged separation from the Baroncelli polyptych [1]. The cusp was never studied by means of analytical techniques before the San Diego Museum gave our team the opportunity when the panel come back to Italy in the context of the exhibition *Giotto, l’Italia* held at the Palazzo Reale of Milano from September 2015 to January 2016. During the Milan exhibition, it was positioned just in front of the polyptych, 2 m high, to suggest to the public the original structure of the polyptych itself (Fig. 6.1).



Fig. 6.1 A possible reconstruction of Baroncelli polyptych, painted by Giotto in the early XIV cent.; both pieces were put in the same room during the exhibit

The exhibition gave the opportunity to investigate the organizational model of Giotto's workshop through analyses of the "Cusp" of San Diego [1].

Because of the impossibility to move the opera during the exhibition period and in according to the exhibition closing hours, our team decided to set up a synergic and complementary protocol of analytical techniques. Whole campaign consisted of 5 diagnostic operations with different operators, 5 h of work each on the close day of the exhibition. The main novelty of this research has been the use of elemental analysis traditionally considered punctual analyses in a mapping way and putting the results in strictly connection with imaging methods. In this way the analytical results revealed a painting technique already used by Giotto [2]. Another important goal reached in this on-the-field research is the use of a new spherical scan for high resolution long-wave scanning in reflectography infrared. In the spherical scanner The motorized head was built with the purpose of allowing the refocusing adjustment needed to compensate the variable camera-painting distance during the rotation of the camera [3]. In this way we get knowledge about the stratigraphy of pictorial material and the underdrawing study: we started from the imaging scans and the visible image; the results achieved have been used to determine the measure points of both punctual analyses such as reflectance spectroscopy and the area for study the effectiveness of scanning portable-XRF (pXRF) [4].

Since many years the use of reflectance spectroscopy and of elemental analysis has been considered a useful tool to investigate pictorial layers without sampling being the first working in visible-near infrared region (380–1050 nm) of the electromagnetic spectrum and so with very poor penetration in pictorial layers and the second coming from X-ray fluorescence, indicatively with higher penetration efficiency in this kind of materials. These two integrated techniques are often considered acting as stratigraphy without sampling from cultural heritage scientists. Now the use with long-wave reflectography and digital false colour infrared could help in an even more accurate type of stratigraphy as long-wave infrared can reach deeper layers inside the painting [5].

6.2 Materials and Methods

6.2.1 *Imaging Techniques*

Long-wave IR reflectography (1000–1700 nm) have been performed with a portable scanning prototype based on a peculiar spherical scanning system [3]. An InGaAs camera is mounted on a motorized three-axes head that allows precise and small movements: 0.01° resolution for rotations and 0.1 mm for the linear stage. Each image has a 256 pixels/in. (10 points per mm) resolution.

The Infrared False Colour (IRFC) imaging analysis have been performed using a DSLR full frame Nikon D800E camera, modified to extend its sensibility from UV to NIR (350–1000 nm). These images obtained in short wave infrared (800–1000 nm)

have a very high resolution (700 pixel/in.; 28 pixel/mm) which is suitable for pigments identification, cracks studies and low hidden layer painting [6].

6.2.2 Analytical Techniques, XRF and FORS

An ELIO spectrometer (XGLab srl, Italy) has been used for elemental analysis due to its portability and to the possibility to obtain an elemental mapping. The instrument is an energy dispersive device XRF (EDXRF, Rh anode, 50 kV, 80 mA, 120 s int. time). We analysed thirty single points to find out the elemental (medium-high Z) composition of the pigments. To get statistical information on the elemental distribution and any possible correlation among them, we performed mapping measurements (1 s for each point) in a specific region, localized around the eyes of God. The system achieved the total XRF map, $60 \times 30 \text{ mm}^2$ (step of $1 \times 1 \text{ mm}$), in about 50 min [4, 7]. ELIO data furnished a three dimensional map of a two dimensional region around the aforesaid region of painting, where the spatial coordinate are completed by a third one which gives the energy of the X-ray detected by the spectrometer. Considering the spectra set as a whole, we defined the Global Spectrum (GS): in this new spectrum every meaningful signal, detected at a given energy, substitutes the intensity, which usually is depicted against energy. This makes the GS a summary of the main X-ray energies in the scanned area, where the position of the local maxima reveals the presence of specific emission elements and its abundance.

On every EDXRF single points of analysis our portable Vis-NIR spectrophotometer (HR4000, Ocean Optics Dunedin FL-US), connected to a stabilized halogen light source, gave us reflectance spectra in the 380–1000 nm (Fig. 6.2). The use of a quartz fibre optics bundle 1.5-m-long, allowed to obtain the reflectance spectra in remote mode, through $45^\circ \times 45^\circ$ measuring geometry. This geometry has been proved to avoid specular reflectance [8, 9] giving results comparable to the one of SPEX by using an integrating sphere probe [10]. We calibrate the spectrophotometer by using Spectralon® 99% and dark trap as white and black reflectance standards respectively. Despite theoretical spectral resolution of 0.27 nm, spectra show a wavelength resolution of about 2.7 nm due to the grating and slit used in this experimental setup. The integration of elemental analyses by XRF and optical reflectance spectra could help in defining stratigraphy without any sampling, due to different penetration depth of different radiation in the pictorial layers [5, 11, 12].

6.3 Results

The reflectography obtained by composing over 691 single shots, $2 \times 3 \text{ cm}$ each (256 pixel/in.) shows the Giotto's characteristic drawing with the use of both thin and large brushstrokes with carbon black ink [6] (Fig. 6.3). Lines of underdrawings are thin and marked for God's hair and beard, while to obtain the saturation he needed on the shady



Fig. 6.2 Probe for reflectance measurements with coaxial contactless fibers optics bundle

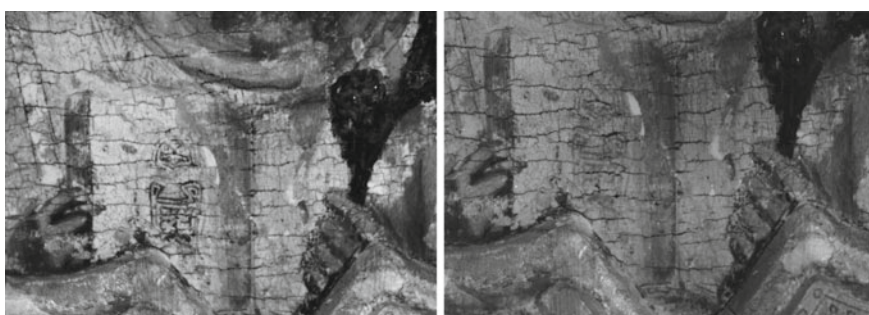


Fig. 6.3 Comparison of short (left) and long-wave (right) IR reflectography: detail of God's hands, high resolution image (700 pixel/in.) shows the Giotto characteristic drawing made by carbon black ink with thin and large brushstrokes

areas, as for the angels' faces, Giotto overlaid large brushstrokes. The master used soft and fluid lines to delineate the angels' hair. There is a question still open about this masterpiece: could the drawing in the cusp be comparable with other drawings and so obtained with the use of patrones? In fact, even if patrones were often used in "fresco" paintings, this technique has been proved widely also in easel paintings. The post-processing images help us to answer to the previous question. The preparatory underdrawing has been extracted by performing an automatic image segmentation on the IR image acquired with the Nikon camera (we chose segmentation threshold = 0.2 taking into account intensity values of darker brushstrokes).

The result has been compared to the IR reflectography image of the face of the blessing Christ of the *Ognissanti Cross* [13]: Fig. 6.4 shows the rigid overlapping of the two images.



Fig. 6.4 Overlapping of underdrawings of God's face (black areas) obtained after image processing on the face of blessing Christ of the Ognissanti Cross

The same algorithm has been applied to the IR images relative to the Angels' faces and the preparatory drawing has been individuated: in this case, a rigid co-registration of the edges of faces, by applying a minimum rotation angle of 0.9° and a maximum rotation angle of 5.8° , has been performed to individuate the possible use of sketches for the drawing of the Angels figures.

Painting materials and drawing techniques of medieval age are similar and well known in the Italian environment [14]. The paintings were realized with egg yolk tempera medium and applied on wood panels. These boards were covered with a mixture of animal glue and chalk (*preparazione*) and a gold foil was often used as back-ground. Nevertheless, scientific examination reveals considerable differences in pictorial techniques.

Giotto was the leading painter in Florence at the beginning of XIV century, and it is stated he changed the course of painting. Cennini [15], the most important biographer of Italian ancient master painters, said that Giotto marked a shift the language of painting from Greek to Latin, meaning that he forsook the Byzantine iconic style for a more classical and naturalistic style.

Starting from data obtained both trough elemental analysis and reflectance spectra, we were able to discover which palette of pigments he used (Table 6.1).

Table 6.1 Giotto's palette and restorations pigments

Colour	Pigments
Blue/light blue	Ultramarine (in some cases with white lead)
	Cobalt blue over white lead (restauration areas)
Green	Chromium oxide green (in some cases with lead/tin based yellow, restauration areas)
Brown	Organic pigment with white lead
	Ochre and vermilion in mixture on a white lead ground
	Vermillion on gold
Red	Vermillion (in some cases with white lead)
	Vermillion on gold
Yellow	Ochre
	Lead/Tin based yellow (restauration areas)
White	White lead
Black	Charcoal black mixed with metal-gall black ink

Original ultramarine and Cobalt blue in the restored areas were found for blue details; for the green areas, Chromium oxide green was found, proving that none of them still show the original pigment; brown was mainly obtained through a mixture of ochre and vermilion; red ochre and vermilion were used for red details; yellow ochre and lead based yellow mixtures for yellow areas, white lead for white ones.

As an example of the joint participation of the techniques applied to the panel, the high resolution image in the visible wavelength range (resolution 700 ppi) is reported in Fig. 6.5a. The false colour image (Fig. 6.5b), with the typical red response due to presence of the ultramarine (as already told, the green areas are from some modern restorations), confirms the punctual analyses performed on the same area.

In fact, the FORS spectrum (Fig. 6.5c) shows the typical trend of ultramarine mixed to white lead (the relative reflectance grows up after 700 nm) and the EDXRF analysis detects, in addition to trace elements as S, Ca, Sr, Fe, Ti, Mn, the presence of Si and Pb (Fig. 6.5d).

Interesting is the black colour present in the book: EDXRF analysis and FORS do not show any characteristic elements or trend. IR reflectography helps in the interpretation of punctual data: the sign of the writing is present in the images acquired with the full frame DSLR Nikon D800E (Silicon detector, Fig. 6.3 left), but the same sign is very weak in the images acquired with the IR scanner (InGaAs detector, Fig. 6.3 right). This is probably due to the presence of a mixture of charcoal black detectable by both the detectors with black metal-gall ink, this last being detectable only by the Silicon detector.

Being the original pigments those typical of Italian late medieval period, it is the structure of the preparation layers that sounds more interesting. By processing the 1800 XRF spectra scanned over the God's face, we were able to obtain a mapping

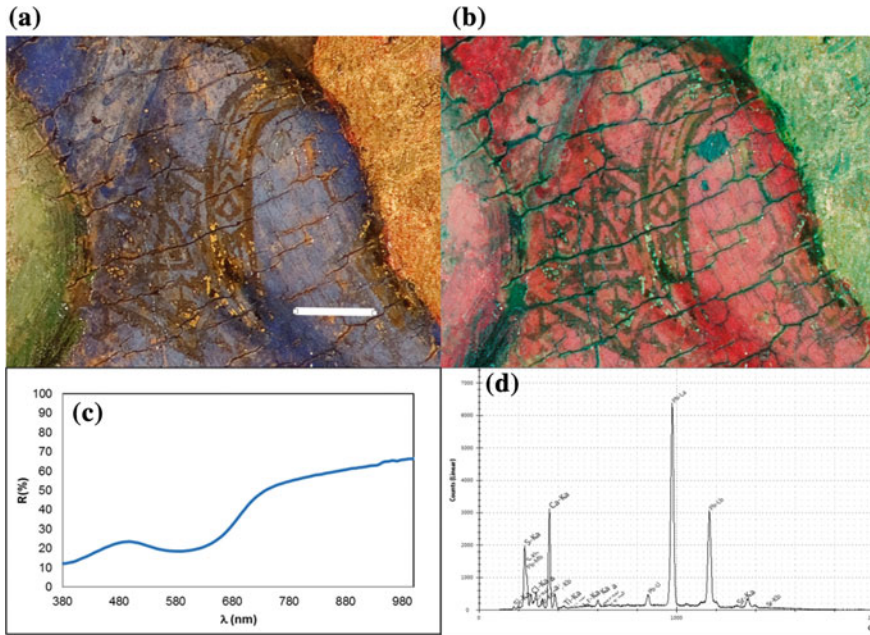


Fig. 6.5 Detail of an angel's shoulder: **a** visible macro, **b** infrared false color, **c** FORS spectrum, **d** XRF spectrum. Analysis reveals the use of ultramarine mixed with white lead

of the most abundant elements. Beside this spatial information, an algorithm that automatically identifies the transition bands from the distribution of the peaks location has been applied [4]: it organizes the spectra as a 3D matrix and integrating into energy it computes the band images. By arranging the band images as vectors and by computing the normalized Cross-Correlation Functions (n-CCFs) the analysis of the matrix of the nCCFs' 0-lag elements is possible. Two groups of band images are closely cross-correlated (the M series and L series of Hg–Pb and the K series of Ca–Fe–K–Sr) but anti-correlated each other. The detection of the M lines of Hg and Pb is probably due to a superficial layer of vermilion and white lead superimposed on a *verdaccio* layer. Verdaccio is a pigment based on a green earth typically used in the Giotto's period. It is interesting to note that this technique was already adopted by Giotto in the S. Giovanni flesh in the Santa Maria Novella Cross [3] as alternative to the monolayer mixture of vermilion, white lead and *verdaccio* implemented in the Christ flesh in the Ognissanti Cross [12]. Interestingly Chromium and Titanium both typical of modern pigments (green and white respectively) does not show significant correlation each other or with other elements (Fig. 6.6). This is imputable to the modern restoration confirmed also by IR false colour image (Fig. 6.7). In fact the XRF images of Cr and Ti slightly differs from the IRFC image due to the XRF sensitivity for short measuring times, 1 s each point. So it gives less accurate information than HR IRFC images.



Fig. 6.6 Distribution maps of Ti and Cr (panel (a) and (b) respectively), indicating the restauration regions highlighted in panel (c)



Fig. 6.7 IR false colour image: the XRF image of Cr and Ti (highlighted in red) slightly differs from the IRFC image (clearer in the picture) because the XRF sensitivity for short measuring times

6.4 Conclusions

We started our research from the analyses of a masterpiece in exhibition, which has been a full-fledged challenge. In fact we had no possibility to move it and only few hours per week to analyze the panel.

Portable, fast and non-invasive synergic techniques, such as IR camera, IR false color, XRF scanner and FORS joined to XRF punctual investigations helped us to reach the aim. These allowed a complete characterization of the pigment palette used by Giotto. Furthermore, thanks to different penetration depths of these techniques and the co-occurrence and correlation of elements in the XRF mapping, we obtained reliable information about the stratigraphic sequence of the painting layers. They suggested us the structure composed by a cinnabar-white lead layer superimposed to a gypsum-verdaccio one, such as in the Cross of Santa Maria Novella. Restoration areas were clearly identified due to the presence of modern pigments, casting light on the conservative history of the cusp. This last is partially unknown starting from the mutilation of the polyptych at the end of the fifteen century till the donation to S. Diego Museum. An unpredictable result of our research was given by imaging technique and post processing elaboration that support the hypothesis of a detailed underdrawing, characterized by wider brush signs. Through image segmentation and pattern recognition algorithms applied to the reflectography, we found many elements to support the thesis about use of patrones for the face of “God the Father”.

Acknowledgements We are grateful to XGLab s.r.l. for kindly placing ELIO MA-XRF scanning spectrometer at disposal. The authors would like to especially thank Michael A. Brown, Associate Curator of European Art at the San Diego Museum of Art.

References

1. S. Romano, P. Petrarola, *Giotto, l'Italia. Catalogo della mostra* (Mondadori Electa, Milano, 2015)
2. M. Ciatti, M. Seidel, *Giotto: the crucifix in Santa Maria Novella*, ed. by M. Ciatti, M. Seidel Paperback (Deutscher Kunstverlag, 2005)
3. M. Gargano, F. Cavaliere, D. Viganò, A. Galli, N. Ludwig, A new spherical scanning system for infrared reflectography of paintings. *Infrared Phys. Technol.* **81**, 128–136 (2017). <https://doi.org/10.1016/j.infrared.2016.12.011>
4. A. Galli, M. Caccia, R. Alberti, L. Bonizzoni, N. Aresi, T. Frizzi, L. Bombelli, M. Gironda, M. Martini, Discovering the material palette of the artist: a p-XRF stratigraphic study of the Giotto panel 'God the Father with Angels'. *X-Ray Spectr.* **46**, 435–441 (2017). <https://doi.org/10.1002/xrs.2751>
5. M. Bertucci, L. Bonizzoni, N. Ludwig, M. Milazzo, A new model for x-ray fluorescence autoabsorption analysis of pigment layers. *X-Ray Spectr.* **39**(2), 135–141 (2010)
6. M. Gargano, N. Ludwig, G. Poldi, A new methodology for comparing IR reflectographic systems. *Infrared Phys. Technol.* **49**, 249–253 (2007). <https://doi.org/10.1016/j.infrared.2006.06.013>
7. M. Alfeld, J. Vaz Pedroso, M. van E., Hommes, G. V. der Snickt, G. Tauber, J. Blaas, M. Haschke, K. Erler, J. Dik, K. Janssens, A mobile instrument for in situ scanning macro-XRF investigation of historical paintings. *J. Anal. Atom. Spectr.* **28**, 760–767 (2013). <https://doi.org/10.1039/c3ja30341a>
8. M. Gargano, N. Ludwig, V. Federighi, R. Sykes, G. Lodi, A. Sardella, A. Carrassi, E.M. Varoni, Fiber optics reflectance spectroscopy (45°x: 45°) for color analysis of dental composite. *Am. J. Dent.* **29**, 223–228 (2016)
9. M. Bacci, A. Casini, C. Cucci, M. Picollo, B. Radicati, M. Vervat, Non-invasive spectroscopic measurements on the Il ritratto della figliastra by Giovanni Fattori: identification of pigments and colourimetric analysis. *J. Cult. Herit.* **4**, 329–336 (2003). <https://doi.org/10.1016/j.culher.2003.09.003>
10. M. Gargano, N. Ludwig, D. Pandini, Use of optical fibre in spectrometry and colorimetry with remote probes. *JAIC J. Int. Colour Assoc.* **8** (2012)
11. L. Bonizzoni, C. Colombo, S. Ferrati, M. Gargano, M. Greco, N. Ludwig, M. Realini, A critical analysis of the application of EDXRF spectrometry on complex stratigraphies. *X-Ray Spectr.* **40**, 247–253 (2011). <https://doi.org/10.1002/xrs.1320>
12. L. Bonizzoni, S. Bruni, A. Galli, M. Gargano, V. Guglielmi, N. Ludwig, L. Lodi, M. Martini, Non-invasive in situ analytical techniques working in synergy: the application on graduals held in the Certosa di Pavia. *Microchem. J.* **126**, 172–180 (2016). <https://doi.org/10.1016/j.microc.2015.12.001>
13. M. Ciatti, *L'officina di Giotto - Il restauro della Croce di Ognissanti* (EDIFIR, Firenze, 2010)
14. L. Bonizzoni, M. Gargano, N. Ludwig, M. Martini, A. Galli, Looking for common fingerprints in Leonardo's pupils using nondestructive pigment characterization. *Appl. Spectrosc.* **71**, 1915–1926 (2017). <https://doi.org/10.1177/0003702817694901>
15. C.D. Cennini, *The Craftsman's Handbook* (Dover Publications, York, 1954)

Chapter 7

Quantitative Analysis of Gold Nano-aggregates by Combining Electron and Probe Microscopy Techniques



Francesca Borghi, Matteo Mirigliano, Paolo Milani and Alessandro Podestà

Abstract The structural and functional properties of nanostructured systems fabricated by the bottom-up assembling of clusters and nanoparticles depend upon the size and shape distribution of the building blocks. Moreover, it is well known that the size and the shape of nano-aggregates can influence toxicity or biological responses. Consequently, analytical techniques able to characterize nano-aggregates size distribution and shape are very important technological tools. This paper is the first step in the development of a quantitative approach combining electron and probe microscopy techniques for measuring the heights and lateral dimensions of primeval incident clusters and nano-aggregates on the substrate. In particular, we have analyzed samples of gold clusters deposited by a supersonic beam on silicon substrates for different deposition time in order to describe the first stage of growth of gold cluster-assembled thin film.

7.1 Introduction

The strategic importance of nanostructured materials relies on the fact that their structural, electronic, magnetic, catalytic, and optical properties can be tuned and controlled by a careful choice and assembling of their nanoscale elemental building blocks [1–4]. The deposition of preformed clusters (aggregations of a few atoms to a few thousands of atoms) as building blocks on a substrate offers the possibility

F. Borghi (✉) · M. Mirigliano · P. Milani · A. Podestà
CIMaINa and Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano,
Via Celoria 16, 20133 Milan, Italy
e-mail: francesca.borghi@unimi.it

M. Mirigliano
e-mail: matteo.mirigliano@unimi.it

P. Milani
e-mail: Paolo.Milani@mi.infn.it

A. Podestà
e-mail: alessandro.podesta@mi.infn.it

© Springer Nature Switzerland AG 2018
P. F. Bortignon et al. (eds.), *Toward a Science Campus in Milan*,
https://doi.org/10.1007/978-3-030-01629-6_7

to carefully control building blocks dimensions and hence to tune the structural and functional properties of the resulting systems [5].

Clusters size distribution, their agglomeration/aggregation state and their shape are only some of the properties which decide for the structural and functional properties of the resulting system [6], crucial in determining the way objects interact and devices work, as in determining the toxicity of nano-aggregates [7]. This in turn requires imaging techniques with a resolution of the order of 1 nm in xy plane and less than 1 Å in the z direction, in addition to the possibility of investigating a wide range of lengths scales. Often several techniques are required to characterize a parameter and minimize the measurement uncertainty.

Atomic Force Microscopy (AFM) imaging makes it possible to measure the height of the objects with sub-nanometric accuracy [5], but the lateral measurements (along x-y axis) contain large errors due to the tip/sample convolution. AFM-images are the results of the interaction between the scanning tip and the sample surface, and this interaction limits the AFM resolution [8–10]. Deconvolution algorithms are currently developed [11–13], but they are relatively difficult to implement and the obtained measurement uncertainties are still large, also because of the continuous change of tip shape and size. One of the alternative method for measuring the lateral dimensions of the objects is Scanning Electron Microscopy (SEM) technique [14], which can measure objects with a resolution of the order of a nanometer. Thus, the combined use of both microscopy techniques enables the characterization of the nano-aggregates in size in the 3-dimensional space.

Today, although imaging techniques are well advanced [15, 16], most images are interpreted qualitatively and only few authors make a quantitative study, which are often based on a manual image-analysis in order to extract morphological information of nano-aggregates and nanostructures. In order to extract quantitative morphological information from AFM/SEM images, we have developed an automated algorithm for sizing nano-aggregates that can easily process many images and that, for each image, can recognize individual objects and calculate the associated morphological parameters and some quantitative shape descriptors. This algorithm is subdivided into two phases: the image pre-processing and the image analysis. The image pre-processing consists in the flattening of the images, in order to get rid of the tilt of the sample and of the AFM scanner bow, and the elimination of the possible bad lines. The image analysis consists into assign a unique label at each object and into measure the morphological properties for each labelled object, i.e.: area, equivalent radius, major and minor axis, height, volume and eccentricity.

In this work, the feasibility of the proposed method has been tested on the dimensional measurements of gold primeval incident clusters (as they arrive on the substrate) deposited by Supersonic Cluster Beam Deposition (SCBD) [17, 18] technique and nano-aggregates grown on the substrate. Heights deduced by AFM and lateral dimensions by SEM images, as the aggregation state of clusters are investigated for samples with different duration of deposition time. Statistical meaning and discrepancy of the morphological parameters calculated from the two imaging methods are compared in order to combine successfully the different informations in all the three dimensions.

7.2 Materials and Methods

7.2.1 Deposition of Gold Clusters by Supersonic Beam

We used the technique of Supersonic Cluster Beam Deposition (SCBD) to deposit gold neutral clusters on a substrate through a supersonic beam [2]. The experimental apparatus, equipped with a pulsed microplasma cluster source (PMCS) [18], is devoted to the production of clusters in gas phase and a subsequent supersonic beam expansion. Briefly, a PMCS consists in a ceramic body with a cavity where a metallic target (Au in the present case) is sputtered. The sputtering process is controlled through a solenoid valve, that faces the cavity. The valve is baked with an inert gas (Ar) at a pressure of about 40 bars and it controls the gas injection opening for hundreds of microseconds. A bias voltage is synchronously applied between the metallic rod and the valve front, in order to produce a discharge that ionizes the gas previously injected, which sputters the rod. The sputtered atoms thermalize with the carrier gas and aggregate in the cavity forming metal clusters. The carrier gas-clusters mixture expands in a low pressure (10^{-7} torr) chamber (the expansion one) through a nozzle out of the PMCS. In the expansion chamber, an aerodynamic focusing nozzle produces a highly-collimated beam [19]. The central part of the beam is selected by a skimmer and then it reaches a second vacuum chamber (the deposition one, at a pressure $10^{-5} - 10^{-6}$ torr), where it impinges on the substrate, supported by a motorized sample holder. The quantity of nano-aggregates that is deposited per unit area and per second (the rate F , measured in $\text{\AA}/\text{s}$) is monitored by a quartz microbalance positioned on the bottom of the sample holder. We deposited with a duty cycle of the source of 4 Hz and with a gas pulse of 211 μs . In these conditions, we have deposited with low rate (about 3–5 $\text{\AA}/\text{s}$), compared to those a PMCS can reach (about 20–30 $\text{\AA}/\text{s}$). In Table 7.1 we report the main deposition conditions of the gold sample analyzed.

7.2.2 Choice of the Substrate

As the roughness of the substrate can influence the imaging quality and the diffusion/aggregation processes of the primeval incident clusters [20], we have analyzed different kind of substrates in order to identify the most suitable one. We selected three types of glass substrates: microscope slides *Thermo-Scientific*, microscope

Table 7.1 Deposition conditions of the gold samples

	Deposition rate ($\text{\AA}/\text{s}$)	Time of deposition (s)
Sample 1	3–5	2
Sample 2	3–5	4

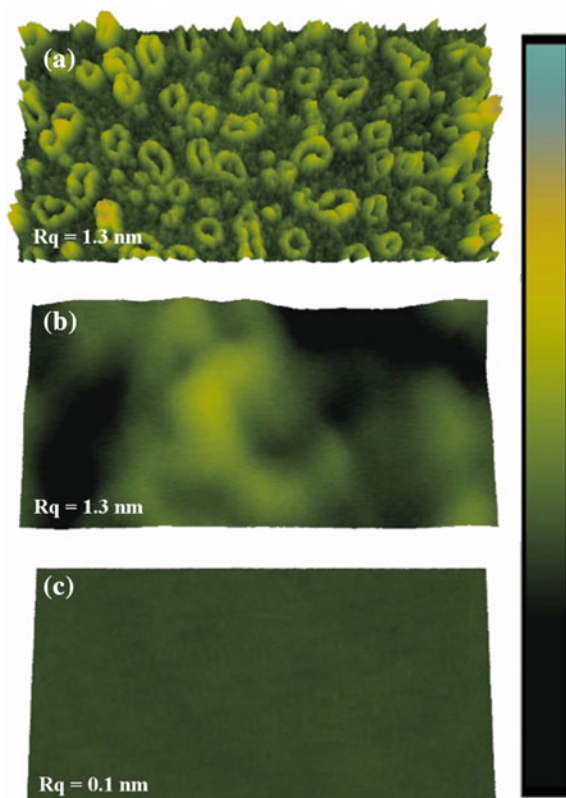


Fig. 7.1 $2\ \mu\text{m} \times 1\ \mu\text{m}$ top-view AFM maps of microscope slides *Thermo-Scientific* (a), microscope slides *Agar Scientific* (b) and a silicon wafer (c). Z-scale ranges from -5 to 10 nm (dark to bright, in the colorbar)

slides *Agar Scientific* and a silicon wafer (*Siegert wafer single side polished*). Before imaging, they were washed in a nitric acid hydrochloride bath and then in ethanol one. The RMS roughness (R_q , standard deviation of the surface heights) of the substrates, characterized by the AFM, is larger for the two microscopies slides ($R_q = 1.3$ nm) than for the silicon wafer ($R_q = 0.08$ nm), as it is also shown in Fig. 7.1. This discrepancy is probably due to the polishing procedure during fabrication processes. Similar results concerning substrate roughness have been found by using only ethanol in cleaning procedure; this excludes the possibility of damages caused by the nitric acid hydrochloride.

We decided to use fragments of silicon wafer as substrate, in order not to influence the growth of the cluster-assembled film, even in sub-monolayer regime, and the evaluation of the nano-aggregates height.

7.2.3 *Electron and Scanning Probe Microscopy*

The investigation of the morphology of the gold sub-monolayer samples was carried out in air using a Multimode AFM equipped with a Nanoscope IV controller (Veeco Instruments). The AFM was operated in Tapping Mode using single crystal silicon tips with nominal radius of curvature of 5–10 nm and cantilever resonance frequency in the range 200–300 kHz. Scan areas were typically $2\ \mu\text{m} \times 1\ \mu\text{m}$ with scan rates of 1–2 Hz. Sampling resolution was typically 1 nm/pixel.

The SEM images have been performed with Zeiss (Supra 40) instrument. We performed images with electron beam of energy 10 keV and on a scanned area of $2.4\ \mu\text{m} \times 1.8\ \mu\text{m}$, with a resolution of 1 nm/pixel.

7.2.4 *Image Processing*

Any image processing protocol for quantitative nano-aggregates sizing is typically divided into two main steps: the image pre-processing, aiming at removing from the images by suitable flattening algorithms distortions introduced during the imaging process, and the main image processing, aiming at extracting quantitative metrological data representative of the system under investigation.

Raw AFM topographs are typically not suitable for a direct analysis, because of several artefacts related to the sample/piezo/cantilever arrangement. Identification and removal of these distortions is an important step when performing accurate quantitative analysis.

AFM topographs typically possess a tilted or curved background. The tilt is typically due to the overall inclination of the sample, while the curvature (“bow”) is due to non-straight motion of the scanner (a quadratic bow is typical of tubular scanner; stacked XY piezos may lead in turn to higher-order distortion). In addition to this global background, the image may be affected by line-by-line offsets; these vertical shifts between adjacent lines may be due to thermal drifts of the z-scanner and/or to thermally-induced changes of cantilever deflection. All these artefacts have an impact on the heights distribution of the image. The flattening procedure is aimed at removing these artefacts, making it possible to quantitatively analyze the three-dimensional AFM images. Figure 7.2 shows a comparison between raw and flattened topographs of Sample 1. The raw image (Fig. 7.2a) shows both an overall tilt due to sample inclination, and line-by-line offsets. Line-by-line (1-dimensional) flattening, consisting in this case in the subtraction of a 1st-order polynomial from each topographic profile (Fig. 7.2b), led to a properly flattened image. If the baseline to be subtracted has been fitted only in the reference flat portion of the image, corresponding to the substrate, we have the best flattened image (Fig. 7.2c) for the characterization of the heights distribution.

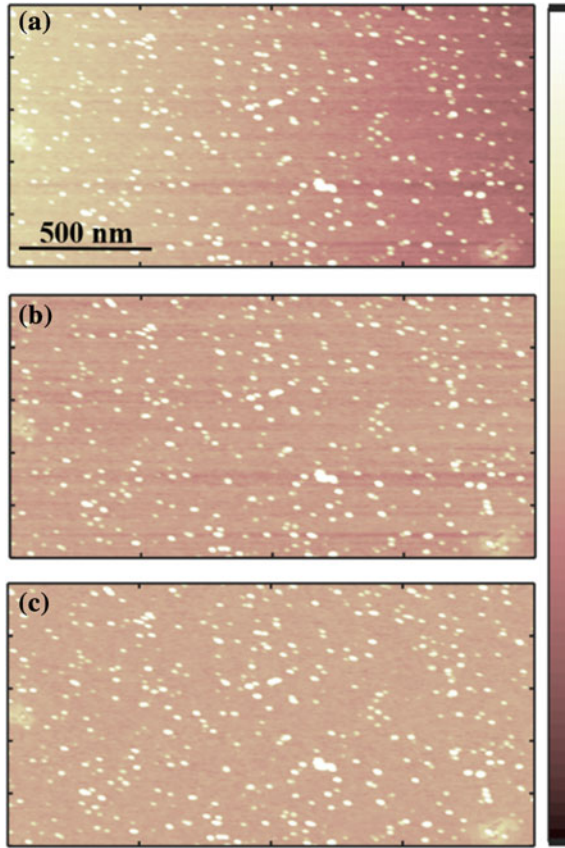


Fig. 7.2 Raw image of Sample 1 (a), flattened image with 1st-order polynomial from each profile (b), flattened image where the baseline to be subtracted has been fitted only in the reference flat portion of the image (c). Z-scale ranges from -2 to 5 nm (dark to bright, in the colorbar)

7.2.5 Data Analysis: Calculation of Geometrical Parameters

In order to calculate morphological parameters for each object in the AFM images, the latter must be recognized and labelled. Objects must be identified and separate from the background; to this purpose, a binary mask identifying the substrate must be created by applying a suitable thresholding method to the flattened AFM topographs. The portion of the image identified by the mask is used to calculate the average quota of the background and its standard deviation σ (i.e. the substrate roughness). The estimation of sample background is an important step in order to calculate the true height value of nano-aggregates (calculated with respect to the substrate) and to calculate a new threshold that identifies the objects of interest.

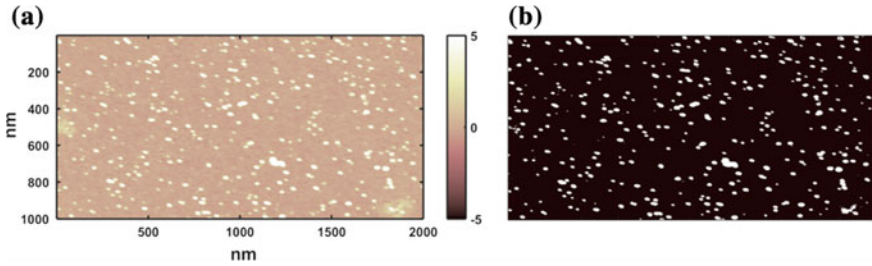


Fig. 7.3 AFM topographical map of Sample 1 (a), with a mask at a z-threshold of $\sim 2\sigma$ (b). The vertical range in (a) is in $[-5, +5]$ nm

We consider “good objects” only those topographic features exceeding by a pre-defined number of standard deviations N , the mean value of the background (Fig. 7.3b). Noticeably, this particle recognition strategy is rather independent on the value of the threshold that is used to build the original substrate mask, because this parameter has a minor influence on the actual value of the mean height value of the substrate and its roughness. Since an image contains multiple objects of interest we must assign a unique label at each object (labelling procedure). Subsequently, we measure the morphological quantities for each labelled object (others can be measured if necessary), among which:

- The height, i.e. the difference between the highest point of the object and the mean value of the background in nm.
- Eccentricity, i.e. a scalar parameter calculated as $\sqrt{1 - \frac{\text{MinorAxis}}{\text{MajorAxis}}}$. This value is between 0 (for a circle) and 1 (for a line segment).
- Major Axis, i.e. the length (in nm) of the major axis of the ellipse that has the same normalized second central moments as the region.
- Minor Axis, i.e. the length (in nm) of the minor axis of the ellipse that has the same normalized second central moments as the region. In particular this parameter is calculated as $\frac{\text{Area}}{(\text{MajorAxis})\pi}$.

In order to analyze the SEM images, we developed an alternative software in Matlab environment devoted to the binarization of the image where the objects are individuated only for the different intensity of the revealed emitted electrons. A thresholding method was chosen to perform a segmentation of the images, because in SEM images the contrast of the substrate is too low in order to be used for threshold identification as in AFM module analysis. This is the most critical step because the subsequent analysis depends on the results of the binarization; in particular, the major problem is the detection of the edges of the nano-aggregates. In fact, on the surface the pixel of the nano-aggregates and the substrate can have similar values and different thresholds can generate different boundary regions causing systematic errors on the measured dimensions. In particular, we used Otsu’s method [21] that automatically perform a thresholding realizing a statistical study on the pixel values of the image. It presents the advantage of being simply reproducible, and it well works with images

where there are isolated nano-aggregates that have high contrast in comparison with the background. All the values below the chosen level are approximated with 0 (the substrate), and the others with 1 (the metal nano-aggregates).

Once the binarization step is carried out, the analysis of the geometrical properties of the objects identified is performed through the same homemade software that exploits Matlab routines used for AFM images analysis.

In addition to the size analysis, we performed a further shape characterization. The nano-aggregates often have stretched or ramified configuration because they are formed from aggregation of primeval clusters. By approximating the cluster with a spherical shape, nano-aggregates can result as a juxtaposition or partial coalescence of metal spherical clusters. To find the region that can be approximated by circles we use a routine that implement the Hough transform [22]. It consists in a voting process in the parameter spaces, where all the image points that can be described by a particular shape are collected and counted in an accumulator variable. The circle parameters that are a local maximum for the calculated accumulators are the best candidate to describe a region [22].

7.3 Results and Discussion

7.3.1 Identification of the Threshold

We investigated the results of the segmentation process through the Otsu's method. It works fine for the recognition of the number of nano-aggregates, thanks to the contrast between them and the background pixels; but it can cause considerable errors in their dimensionality estimation because of the image resolution (1 nm/pixel). In fact, the distribution of the pixels among a large interval of values is the main problem in the calculation of a threshold. This slightly differs from the correct one that identifies the whole nano-aggregates. For this reason we applied a correction factor $\Delta = 0.19$, to take this error into account.

Figure 7.4 shows how the segmentation of the image works through a thresholding method. In particular, it is possible to see how the segmentation of an island changes varying the threshold from the Otsu's one. In images with a resolution of about 1 nm/pixel, the choice of a different threshold can cause the variation of the number of pixels in the recognized object and so an error on the measured dimension of few nanometer. The variation of the threshold value can produce segmentation processes which can alter the number of nano-objects on the substrate. This is the reason why the thresholding step is the most critical one.

Figure 7.5 shows a comparison between the minor and major axis normalized distributions of the nano-aggregates for a SEM image of Sample 1, obtained by using two thresholding methods: the AFM module and Otsu's one. The multi-modal

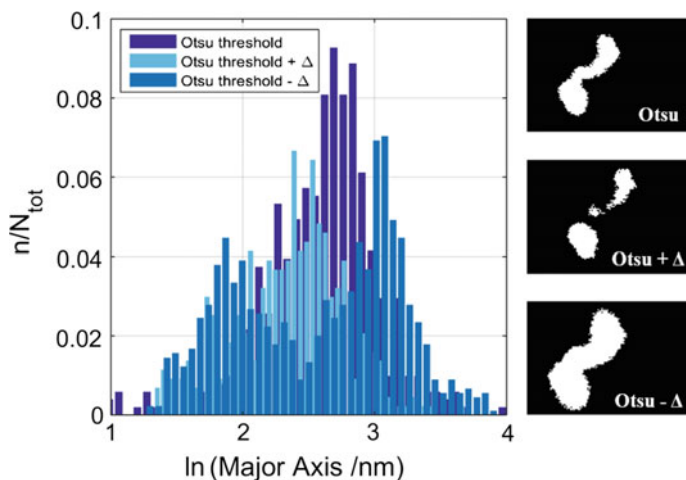


Fig. 7.4 Normalized histogram of the major axis of the gold nano-aggregates identified in Sample 1, with different threshold values. On the right, the corresponding imaging of the nano-objects

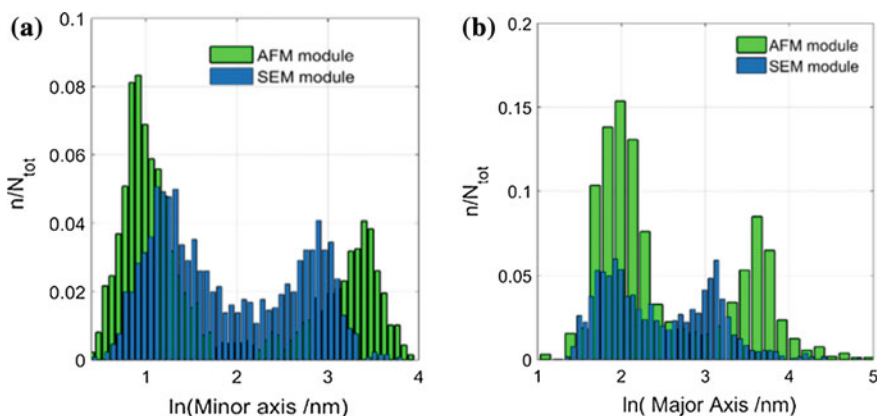


Fig. 7.5 Normalized distribution of minor (a) and major (b) axis of the gold nano-aggregates in Sample 1, analyzed with AFM and SEM module software

distribution is confirmed, while the distance between the two peaks in the distributions is due to the error committed because of the image resolution, which is related to the choice of the threshold as mentioned before. This result suggests an intrinsic uncertainty in the identification of the lateral dimensions of the nano-objects due to the identification of the thresholds in the analysis process.

7.3.2 Analysis of the Size Distribution of Gold Nano-aggregates

In Fig. 7.6 we report two representative AFM top-view images of the evolving sub-monolayer morphology of gold nano-aggregates with increasing quantity of primeval incident clusters deposited. Tip-sample convolution affects only the lateral dimensions of the imaged objects; heights are not affected in the case of this rigid particles, therefore a measurement of the height of the objects, under the hypothesis that the objects are spheroidal, represents an accurate estimation of their diameters.

The normalized distributions of the heights of the objects analyzed in the two samples are reported in semi-log scale [23] in Fig. 7.7. For each system analyzed, the size distributions (heights distribution) are broad and multi-modal. Both the depositions are more populated by very small clusters, with height peaked at 0.4 ± 0.2 nm. We can clearly recognize this population of very small objects also in SEM images, so we can exclude that they come from a silicon substrate contribution. Since the first peak is related to clusters with lateral dimensions close to 1 nm (also in x-y dimensions) in an image with resolution of 1 pixel/nm, we decided to exclude this part of particle population from the further quantitative analysis. Inertial effects of clusters in the supersonic beam [19, 24] determine the concentration of also larger nano-aggregates along the beam axis, with height peaked at 5 ± 2 nm. We can assume that these larger clusters have the prominent role in the cluster-assembled film growth.

There is no evidence of a shift in the heights distribution of the gold nano-aggregates with increasing time of deposition, which suggests that no growth in z-direction has still taken place. Anyway, the counts beneath the Gaussian profiles are double in Sample 2 distribution compared to the one of Sample 1. This trend means that the number of primeval incident clusters deposited is proportional to the time of deposition, and that the further clusters deposited on Sample 2 effectively land on the substrate and not on the top of the pre-deposited clusters.

The plot of the eccentricities (calculated from SEM images, Fig. 7.8) shows the presence of a great fraction of population of nano-aggregates with an elliptical shape. We plot the normalized distribution of the minor and major axes (Fig. 7.9) in semi-

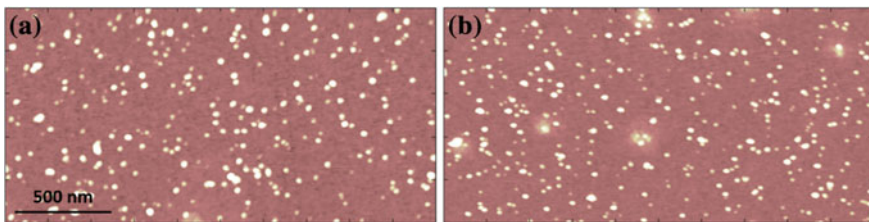


Fig. 7.6 AFM top-view images of Sample 1 (a) and Sample 2 (b). Z-scale ranges from -2 to 5 nm (dark to bright)

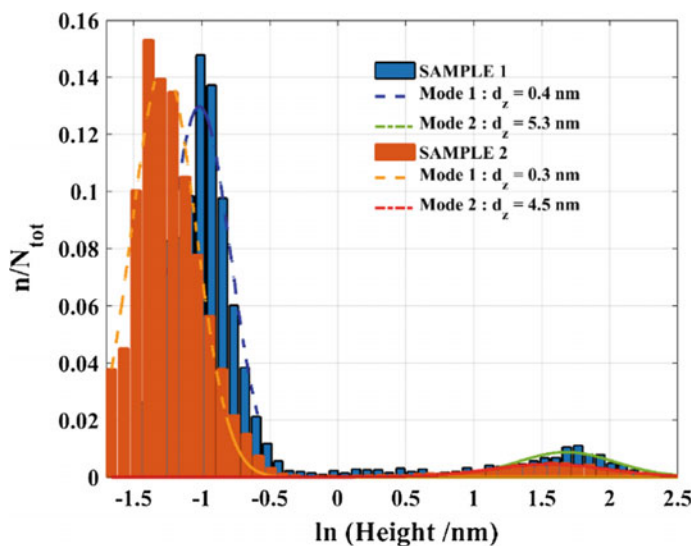


Fig. 7.7 Normalized distribution of primeval incident clusters heights

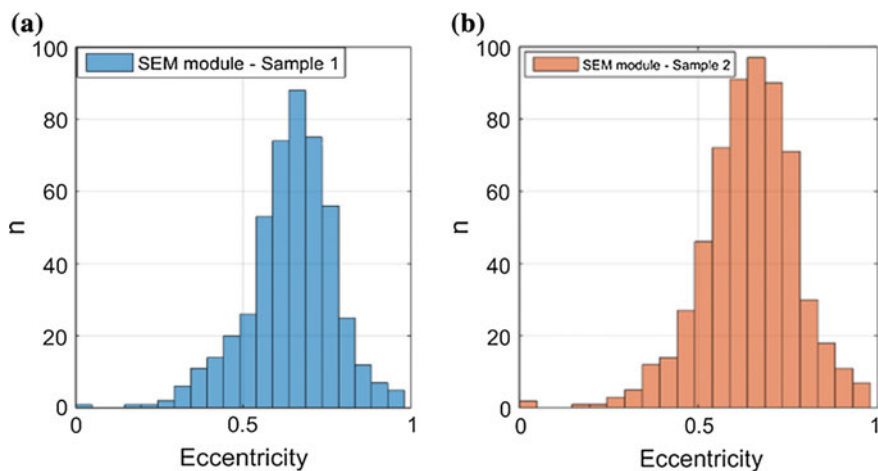


Fig. 7.8 Distribution of eccentricity of gold nano-aggregates in Sample 1 (a) and Sample 2 (b)

log scale, as defined in the Sect. 7.2.5, to characterize the lateral size of the gold nano-aggregates on the substrate for the two samples.

They are multimodal and each mode was approximated to a Gaussian in a semi-log scale. In Sample 1, the two main peaks of the minor axis distribution back-transformed are approximately at 3 and 16 nm, while the major axis around 7 and 20 nm. The multimodal distribution reflects the multi-dispersion of the primeval incident clusters and nano-aggregates size also in z direction. In Sample 2 we can

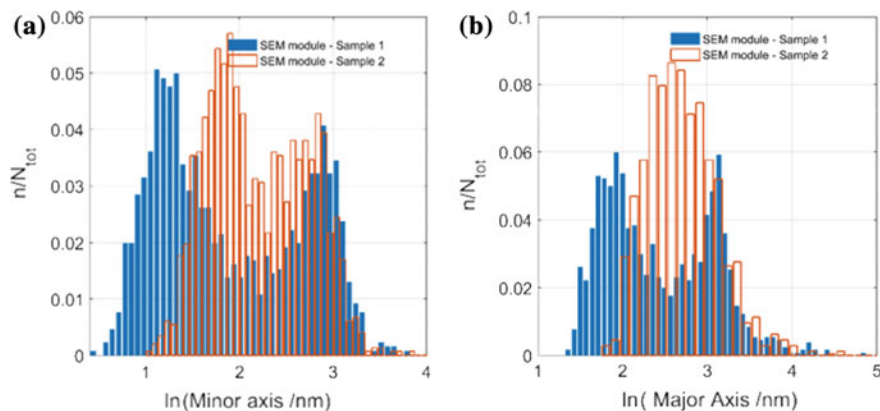


Fig. 7.9 Normalized distribution of minor (a) and major (b) axis of gold nano-aggregates in Sample 1 and Sample 2

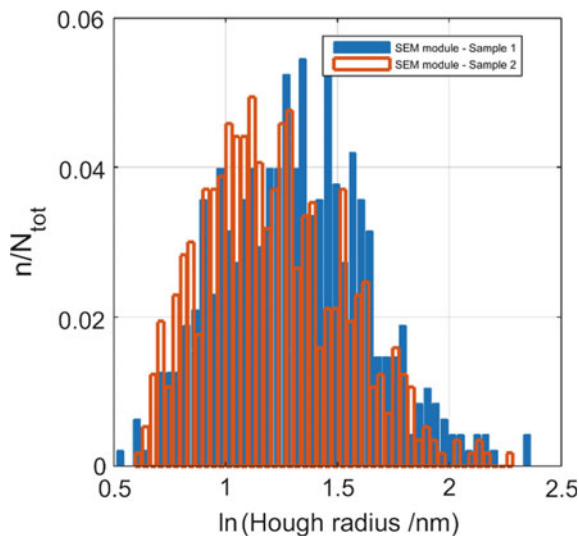


Fig. 7.10 SEM image of Sample 2. On the right, some details of the image with superimposed circle with Hough radius

appreciate a re-distribution of the peaks, in particular an increase in the value of the smaller one. This evidence suggests that laterally the growth of nano-aggregates takes place at this stage by cluster assembling, whatever it is with coalescence or juxtaposition phenomena.

In order to investigate the merging phenomena between clusters we have characterized the Hough radius of the nano-aggregates identified in the SEM images. Figure 7.10 is a SEM image of Sample 2, with superimposed the Hough analysis of some aggregates on the right of the figure.

Fig. 7.11 Normalized distribution of the Hough radius of clusters



The normalized distributions of the Hough radius of the objects of the two samples are shown in Fig. 7.11, in semi-log scale. The peak is around 3.5 nm for both the samples.

The nano-aggregates which can be identified in the early stage of growth of gold cluster-assembled film are composed by units approximately 7 nm large, which are juxtaposed on the substrate.

7.4 Conclusions

We have characterized the first stage of growth of gold cluster-assembled thin film in sub-monolayer regime, with an approach based on the combined AFM and SEM techniques. AFM analysis is the best choice in order to characterize the dimension in z direction of nano-aggregates on a substrate. This technique offers also the possibility to distinguish very small objects (height of fraction of nanometer) which cannot be accurately distinguish in SEM image because of the resolution of the image. Also the identification of a threshold for the characterization of the geometrical properties of the objects, thanks to the standard deviation of the heights of the substrate, is a good strategy to avoid systematic error in the further analysis steps.

SEM technique allows the possibility to explore the lateral dimensions and the growth of clusters with a high statistical importance. Anyway, we have highlighted the limit of this technique, in particular in the analysis process concerning the identification of the best threshold based on the gray-scale.

By combining these results from the analysis of AFM and SEM images we can describe the aggregation state of the primeval incident clusters on the substrate. In

particular, the major mode of the primeval incident clusters is characterized by units of gold clusters of 5 nm in height (by AFM measurements) and 7 nm in width (by SEM analysis). In the early stage of growth, the units that compose a nano-aggregate can be distinguished for their boundaries in SEM images thanks to a different gray-level. In fact, they are juxtaposed or not completely merged into a single larger new nano-object. Lateral growth of nano-aggregates is preferred to growth in z-direction.

Acknowledgements Prof. Michele Perego and Dr. Fabio Zanenga from the MDM Laboratory, IMM-CNR (Via Olivetti 2, 20864, Agrate Brianza, Italy), are gratefully acknowledged for the SEM analysis.

References

1. Y. Huttel, *Gas-Phase Synthesis of Nanoparticles* (Wiley, 2017)
2. K. Wegner, P. Piseri, H.V. Tafreshi, P. Milani, *J. Phys. Appl. Phys.* **39**, R439 (2006)
3. P. Grammatikopoulos, S. Steinhauer, J. Vernieres, V. Singh, M. Sowwan, *Adv. Phys. X* **1**, 81 (2016)
4. C. Binns, *Surf. Sci. Rep.* **44**, 1 (2001)
5. A. Podestà, F. Borghi, M. Indrieri, S. Bovio, C. Piazzoni, P. Milani, *J. Appl. Phys.* **118**, 234309 (2015)
6. F. Borghi, PhD thesis, Università degli Studi di Milano (2015)
7. M.A.C. Potenza, Ž. Krpetić, T. Sanvito, Q. Cai, M. Monopoli, J.M. de Araújo, C. Cella, L. Boselli, V. Castagnola, P. Milani, K.A. Dawson, *Nanoscale* **9**, 2778 (2017)
8. L. Chen, C.L. Cheung, P.D. Ashby, C.M. Lieber, *Nano Lett.* **4**, 1725 (2004)
9. L. Chen, X. Yu, D. Wang, *Ultramicroscopy* **107**, 275 (2007)
10. Y. Ebenstein, E. Nahum, U. Banin, *Nano Lett.* **2**, 945 (2002)
11. J.S. Villarrubia, *Surf. Sci.* **321**, 287 (1994)
12. P. Nagy, G.I. Márk, E. Balázs, *Microbeam Nanobeam Analysis* (Springer, Vienna, 1996), pp. 425–433
13. H. Itoh, T. Fujimoto, S. Ichimura, *Rev. Sci. Instrum.* **77**, 103704 (2006)
14. R. Reichelt, *Microscience* (Springer, New York, NY, 2007), pp. 133–272
15. A. San, Paulo and R. García, *Biophys. J.* **78**, 1599 (2000)
16. C. Möller, M. Allen, V. Elings, A. Engel, D.J. Müller, *Biophys. J.* **77**, 1150 (1999)
17. E. Barborini, P. Piseri, P. Milani, *J. Phys. D Appl. Phys.* **32**, L105 (1999)
18. H.V. Tafreshi, P. Piseri, G. Benedek, P. Milani, *J. Nanosci. Nanotechnol.* **6**, 1140 (2006)
19. H.V. Tafreshi, G. Benedek, P. Piseri, S. Vinati, E. Barborini, P. Milani, *Aerosol Sci. Technol.* **36**, 593 (2002)
20. P. Jensen, *Rev. Mod. Phys.* **71**, 1695 (1999)
21. N. Otsu, I.E.E.E. *Trans. Syst. Man Cybern.* **9**, 62 (1979)
22. H. Yuen, J. Princen, J. Illingworth, J. Kittler, *Image Vis. Comput.* **8**, 71 (1990)
23. E. Limpert, Stahel Werner A., and M. Abbt, *Bioscience* **51**, 341 (2001)
24. E. Barborini, I.N. Kholmanov, A.M. Conti, P. Piseri, S. Vinati, P. Milani, C. Ducati, *Eur. Phys. J. At. Mol. Opt. Phys.* **24**, 277 (2003)

Chapter 8

The Discovery of the Higgs Boson and Most Significant Measurements at the LHC



Marcello Fanti

Abstract On July 4th, 2012, CERN announced the discovery of a new particle, with properties quite similar to those expected from the Standard Model for the Higgs boson. The observation was carried out by the ATLAS and CMS collaborations, analysing the proton-proton collisions produced at the Large Hadron Collider (LHC), at energies of 7 and 8 TeV. More data were collected since then, confirming that indeed the long-sought Higgs boson was finally found. In this document, the main characteristics of this particle are summarized, together with the most significant measurements of its properties, obtained from a combined analysis of the ATLAS and CMS data recorded at 7 and 8 TeV, during the Run-I, in the years 2010–12. The mass of the Higgs boson is $m_H \simeq 125$ GeV. The measured cross-section, couplings, and spin-parity status are compatible with Standard Model expectations. The most recent measurements, based on the Run-II data at a collision energy of 13 TeV, won't be covered in this document, as a combination of the two experiments is not available yet.

8.1 Introduction

The Standard Model (SM) of elementary particle physics is a theory that describes the electroweak (EW) and strong fundamental interactions by means of gauge theories [1]. The matter is made of spin-1/2 “fermions”: quarks (q), charged leptons (ℓ^\pm), and neutrinos (ν). Fermions interact by exchanging spin-1 vector bosons, also called “gauge bosons”, as described by the Feynman diagrams:

M. Fanti (✉)

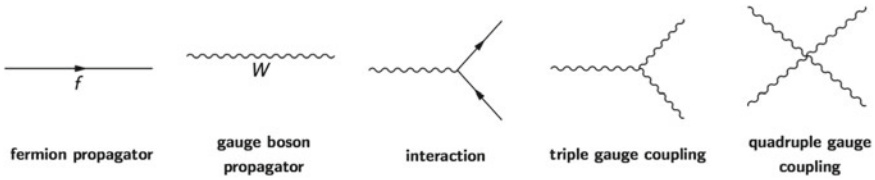
Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy
e-mail: marcello.fanti@mi.infn.it

M. Fanti

INFN, Sezione di Milano, Milan, Italy

© Springer Nature Switzerland AG 2018

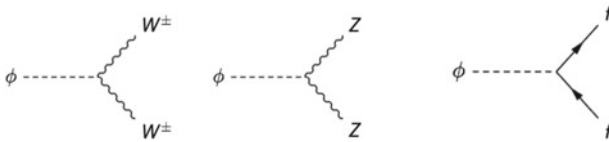
P. F. Bortignon et al. (eds.), *Toward a Science Campus in Milan*,
https://doi.org/10.1007/978-3-030-01629-6_8



Each interaction is described by a Lagrangian, invariant under a symmetry group. The EW interactions are described by an $SU(2)_L \times U(1)_Y$ symmetry group, where the $SU(2)_L$ part treats interactions of left-handed chiral doublets of fermions, while the $U(1)_Y$ part couples to all fermions according to their weak hypercharge. The strong interaction is described by the Quantum Chromodynamics (QCD), with an $SU(3)$ symmetry group acting on colors. The SM predictions are very accurate and verified experimentally to high precision, since the time of previous accelerators like LEP [3].

Gauge theories assume that gauge bosons are massless. This is indeed the case for photons (γ) and gluons (g), which mediate respectively the electromagnetic and strong (QCD) interactions. On the opposite, the mediators of the weak interaction—the W^\pm and Z bosons—have sizable masses, about 80 and 91 times the mass of a proton. However, introducing explicit mass terms in the Lagrangian would spoil the symmetry.

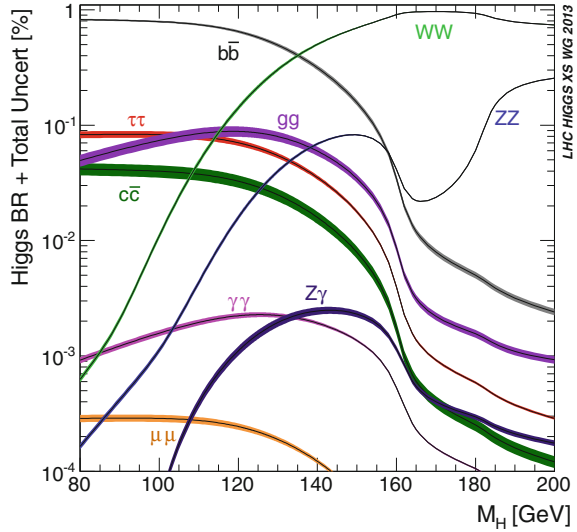
To solve this paradox, the so-called “Higgs mechanism” was proposed [2], back in the 60ies of the past century, by Peter Higgs, and independently by François Englert and Robert Brout. This model postulates the existence of a scalar field ϕ , with a uniform non-vanishing value in vacuum, such that it can interact with most particle fields.



The gauge symmetries are preserved, but the symmetry of the vacuum state is broken. As a consequence, all particles interacting with the scalar field acquire their masses. Since the ϕ -field does not carry electric charge nor color charges, photons and gluons do not interact with it and therefore are left massless.

The suggested experimental test of the model, was to try to excite the ϕ -field, such to materialize its quantum, the “Higgs particle”, or “Higgs boson”. This could be achieved by colliding accelerated particles, with a sufficient energy to materialize the Higgs boson mass, m_H . As this mass is a free parameter in the model, there was no clue as to what energy was necessary. Former particle physics experiments sought the Higgs boson without success. The LHC was the first proton-proton collider to reach an energy sufficient to produce the Higgs boson, which now we know has a mass $m_H \simeq 125$ GeV.

Fig. 8.1 Branching ratios of the Higgs boson into several decay channels, as predicted by SM as a function of the Higgs mass m_H



8.2 Discovery

According to SM, the Higgs boson is foreseen to decay promptly to pairs of lighter particles, in several ways, such as:

$$H \rightarrow b\bar{b}, \quad H \rightarrow \tau^+\tau^-, \quad H \rightarrow W^+W^-, \quad H \rightarrow ZZ, \quad H \rightarrow \gamma\gamma$$

(where b, \bar{b} are the bottom quark and anti-quark, and τ^\pm are the 3rd generation charged leptons). The branching ratios of all decay channels are predicted, as a function of its mass, m_H , as shown in Fig. 8.1.

The easiest decay channels to detect at collider experiments are $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ \rightarrow 4\ell$ —the latter occurring when both Z -bosons decay each in a charged lepton pair, $Z \rightarrow \ell^+\ell^-$.¹ The produced final states are in both cases easy to identify, all involved particles are detectable, and their energy measures well calibrated. Examples of such final states are visible in the event displays shown in Fig. 8.2.

For all such events, the *invariant mass* is computed:

$$m_{\gamma\gamma} = \sqrt{(E_{\gamma 1} + E_{\gamma 2})^2 - |\mathbf{p}_{\gamma 1} + \mathbf{p}_{\gamma 2}|^2}; \quad m_{4\ell} = \sqrt{\left(\sum_{\ell=1}^4 E_\ell\right)^2 - \left|\sum_{\ell=1}^4 \mathbf{p}_\ell\right|^2}$$

¹In the present document, charged leptons (ℓ^\pm), are meant to be electrons (e^\pm) or muons (μ^\pm), which can cross the detector without decaying, and therefore can be identified. The τ -leptons are not considered, as they decay very close to the interaction vertex, making their identification more challenging.

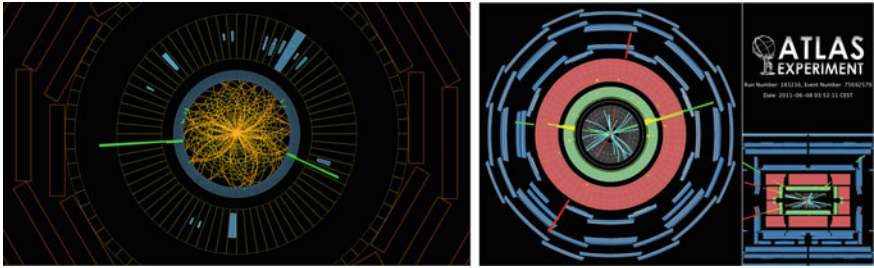


Fig. 8.2 Event displays of Higgs boson candidates, decaying to $\gamma\gamma$ (by CMS, on the left) and to $e^+e^-\mu^+\mu^-$ (by ATLAS, on the right)

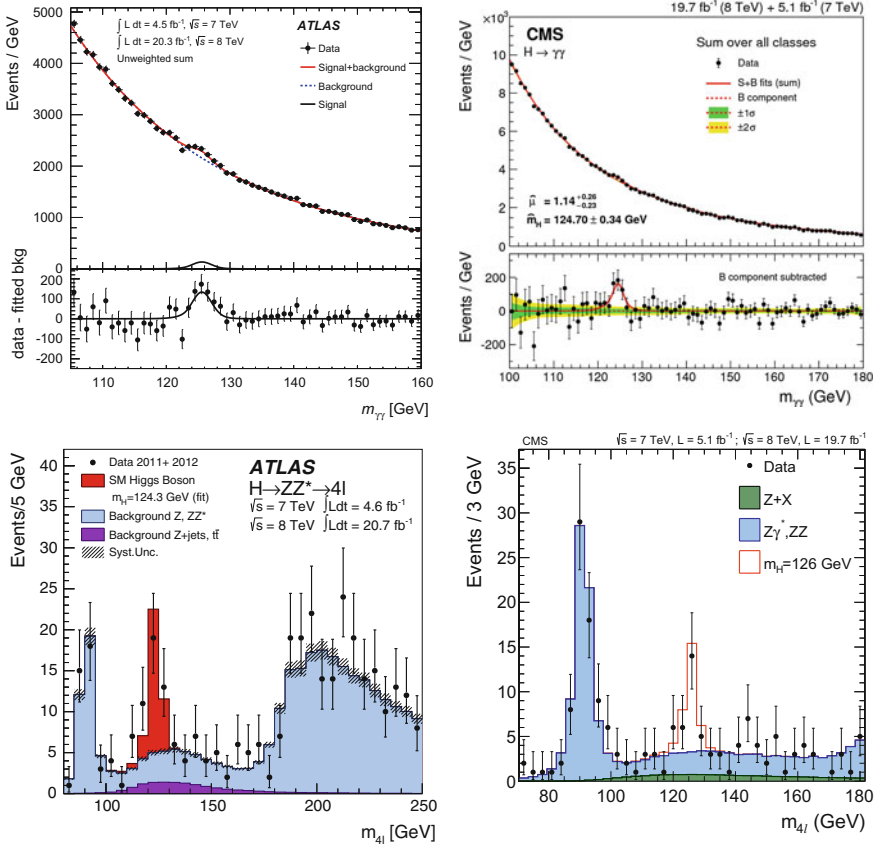


Fig. 8.3 Invariant mass plots, for $\gamma\gamma$ (top plots) and 4ℓ (bottom plots) final states, as recorder by ATLAS (left plots) and CMS (right plots)

where (E, \mathbf{p}) is the 4-momentum of each particle, and the two sums are carried out on the two photons, or the four leptons, respectively.

Several processes contribute to form $\gamma\gamma$ and 4ℓ final states, thus the observable spectra of $m_{\gamma\gamma}$ and $m_{4\ell}$ are generally broad. However, for events arising from a decay of a particle, the invariant mass must be equal to the mass of the decayed particle, within the experimental uncertainty.

Figure 8.3 shows invariant mass plots for $m_{\gamma\gamma}$ and $m_{4\ell}$, as recorded by ATLAS and CMS: in all cases, an excess of events around 125 GeV is observed. These four observations provided the evidence for the discovery of a new particle, as announced by ATLAS [4] and CMS [5] in 2012.

8.3 Measurements of the Higgs Boson Properties

Once the discovery was established, more decay modes were observed: $H \rightarrow W^+W^-$ and $H \rightarrow \tau^+\tau^-$, besides the already mentioned $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ \rightarrow 4\ell$. The next goal was then to investigate all the properties of the new particle, such to ensure whether it is indeed the Higgs boson. In the SM, the only free parameter of the Higgs mechanism is the Higgs boson mass. Once it is determined, the theory becomes fully predictive and can be tested.

8.3.1 Mass Measurement

The invariant mass plots of Fig. 8.3 allow to extract the particle's mass, by combining the four measurements [6], as shown in Fig. 8.4. The best estimate is

$$m_H = (125.09 \pm 0.21_{(\text{stat})} \pm 0.21_{(\text{syst})}) \text{ GeV}$$

This measurement relies on an accurate calibration of the energy measurement, for charged leptons and photons. In the case of the charged leptons, the calibration exploits the observed decays of Z -bosons, $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$, which provide a very good ‘‘candle’’, the Z -mass being known to high precision from LEP measurements. For photons, the calibration is more indirect, being based on the similarity of the energy deposits of photons and electrons in the electromagnetic calorimeters.

8.3.2 Spin, Parity, and Charge Conjugation

According to the SM, the Higgs particle is electrically neutral. Moreover, it has $J^{CP} = 0^{++}$, i.e. it has spin = 0 and even parity and charge-conjugation. The null

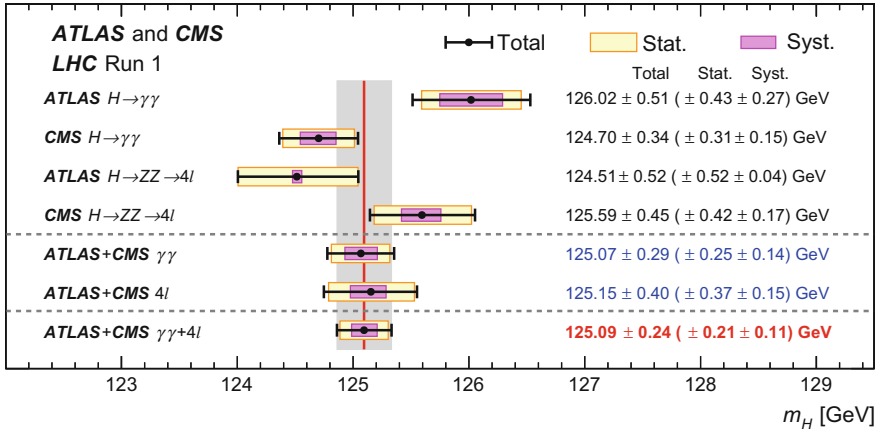


Fig. 8.4 Measurements of the Higgs boson mass, from the $\gamma\gamma$ and the 4ℓ decay modes, as recorded by ATLAS and CMS, and their combinations

electrical charge, as well as the integer spin, can be obviously inferred from any of the observed decay modes. Moreover, the $H \rightarrow \gamma\gamma$ decay proves that $C = +1$ and rules out the spin-1 hypothesis [7], leaving only the spin-0 and spin-2 options. The analysis of the angular distributions of the decay products, expected to be isotropic for spin-0, excludes the spin-2 hypothesis. The analysis of decay angles in the $H \rightarrow ZZ \rightarrow 4\ell$ and $H \rightarrow W^+W^- \rightarrow e\nu_e\mu\nu_\mu$ channels allows to probe the polarizations of the ZZ and W^+W^- gauge bosons, which result being compatible with even parity, $P = +1$ [8].

8.3.3 Cross-Sections and Couplings

The nature of the Higgs mechanism, that dynamically assigns masses to the particles, implies that the couplings of the Higgs boson to fermions (g_f) and to gauge bosons (g_V) must depend on particles' masses, according to:

$$g_f = \frac{m_f}{v}; \quad g_V = \frac{2m_V^2}{v}$$

(where m_f , m_V are the fermion's and gauge boson's mass, and $v \simeq 246$ GeV is the value of the Higgs field in vacuum).

Once m_H is given, SM allows to compute the cross-section of the Higgs boson production, as displayed in Fig. 8.5. Several production modes are possible. The dominant one is the gluon-fusion ("ggF", $gg \rightarrow H$), accounting for about 90% of the total cross-section, in which a gluon pair produce a Higgs particle via a virtual loop containing a top quark or a bottom quark. Other important processes are the

Fig. 8.5 Cross-section of the Higgs boson production, as a function of the collision energy, for several production modes

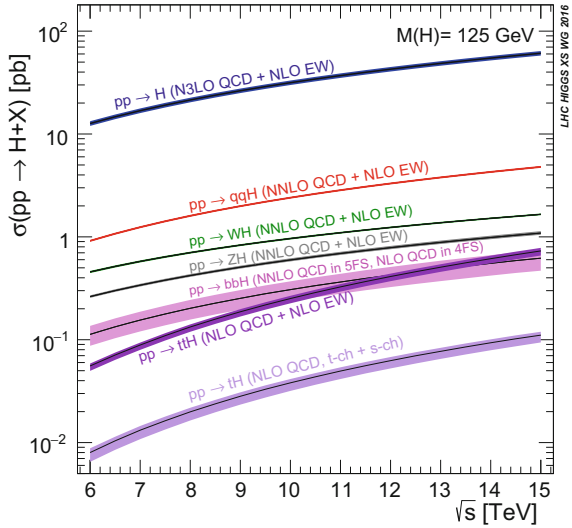
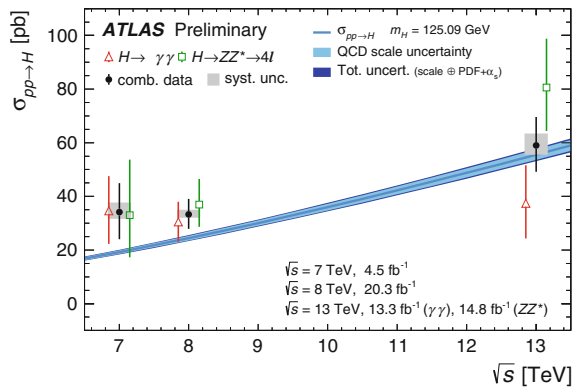


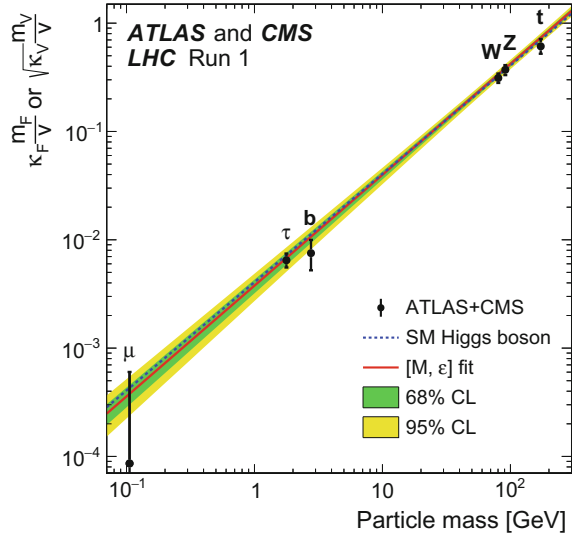
Fig. 8.6 Predicted cross-section of the Higgs boson production (blue band), compared with measured values at several collision energies



vector-boson fusion (“VBF”, $WW, ZZ \rightarrow H$) and the “Higgs-strahlung” (“VH”, $q\bar{q} \rightarrow WH$ and $q\bar{q} \rightarrow ZH$). Both are subdominant, yet they are experimentally identifiable. In the VBF, the WW or ZZ that produce the Higgs boson are radiated off protons’ valence quarks, that as a consequence undergo a significant deflection, such to produce energetic hadronic jets in the forward and backward regions of the detector. In the VH mode, the Higgs decay products are accompanied by the decay products of the associated W or Z —a pair of hadronic jets for the cases $W, Z \rightarrow q\bar{q}$; or a pair of charged leptons for $Z \rightarrow \ell^+\ell^-$; or a charged lepton and some undetected energy due to the neutrino, for $W \rightarrow \ell\nu$.

Combining all possible processes, the total cross-section can be computed, and then compared with the experimental measurements. The result is shown in Fig. 8.6. The measurements have been achieved on the $\gamma\gamma$ and 4ℓ decay channels [9].

Fig. 8.7 Measured couplings to fermions and gauge bosons, as a function of particles' masses. The vertical axis displays g_f for fermions (τ , b , t) and $\sqrt{\frac{g_V}{2v}}$ for gauge bosons (W , Z), such that both are expected to scale as $\frac{m}{v}$ and can be drawn on the same plot



To probe more in detail the g_f , g_V couplings, the rates of all observed decay channels and the main production modes are combined. The result, obtained from a combination of ATLAS and CMS measurements [10], is summarized in Fig. 8.7. This result proves the prediction of the SM over a span of 3 orders of magnitude. The sensitivity to g_b , g_t comes mainly from the ggF process, where the top and bottom quarks enter the virtual loop. The $H \rightarrow \mu^+ \mu^-$ is expected to have a tiny branching ratio, and it has not been observed so far, therefore the corresponding point in the plot is a result of the global fit.

8.4 Conclusions

During the Run-I of LHC, with proton-proton collisions at energies of 7 and 8 TeV, the existence of the Higgs boson was experimentally established, and its properties were measured, confirming the expectations. ATLAS and CMS measurements have been combined to give “world averages”.

The $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ \rightarrow 4\ell$ decay channels allowed to measure its mass, $m_H = (125.09 \pm 0.21_{(\text{stat})} \pm 0.21_{(\text{syst})})$ GeV. The combined analysis of $H \rightarrow \gamma\gamma$, ZZ , WW decay modes allowed to determine the spin-parity status to be $J^{CP} = 0^{++}$, as expected. The total cross-section is compatible with the SM predictions at collision energies of 7, 8, 13 TeV. The combination of all observed decay modes and production modes (gluon-fusion, vector-boson fusion, Higgsstrahlung), allowed to probe the Higgs boson couplings to fermions and gauge bosons.

The results are in perfect agreement with the Standard Model predictions, however the uncertainties are still large.

Since the beginning of Run-II, with a collision energy of 13 TeV, all measurements are being repeated, and ATLAS and CMS results will eventually be combined again. The results from Run-II will contribute to improve the measurements, due to the larger cross-section and collected luminosity. Recently, the observations of the $H \rightarrow b\bar{b}$ decay [11] and of the $t\bar{t}H$ production [12] have been achieved: both will allow a better determination of the g_b , g_t couplings. The choice of focusing this report on Run-I data was driven only by the existence of the ATLAS+CMS combinations. New results at 13 TeV are appearing as soon as ready on the ATLAS and CMS public web pages [13].

References

1. S.L. Glashow, Partial-symmetries of weak interactions. Nucl. Phys. **22**(4), 579 (1961); S. Weinberg, A model of leptons. Phys. Rev. Lett. **19**, 1264 (1967); A. Salam, Weak and electromagnetic interactions, in Proceedings of the Eighth Nobel Symposium in Elementary Particle Theory: Relativistic Groups and Analyticity, N. Svartholm, (ed.) Almqvist & Wiksell, p. 367 (1968); G.T. Hooft, M. Veltman, Regularization and renormalization of gauge fields. Nucl. Phys. **B44**, 189 (1972)
2. F. Englert, R. Brout, Broken symmetry and the mass of gauge vector mesons. Phys. Rev. Lett. **13**, 321 (1964); P.W. Higgs, Broken symmetries, massless particles and gauge fields. Phys. Lett. **12**, 132 (1964); P.W. Higgs, Broken symmetries and the masses of gauge bosons. Phys. Rev. Lett. **13**, 508 (1964); G.S. Guralnik, C.R. Hagen, T.W.B. Kibble, Global conservation laws and massless particles. Phys. Rev. Lett. **13**, 585 (1964); P.W. Higgs, Spontaneous symmetry breakdown without massless bosons. Phys. Rev. **145**, 1156 (1966); T.W.B. Kibble, Symmetry breaking in non-Abelian gauge theories. Phys. Rev. **155**, 1554 (1967)
3. ALEPH, DELPHI, L3, OPAL collaborations, precision electroweak measurements on the Z resonance. Phys. Rep. **427**, 257–454 (2006)
4. ATLAS collaboration, Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. Phys. Lett. B. **1**, 716 (2012)
5. CMS collaboration, Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. Phys. Lett. B. **30**, 716 (2012)
6. ATLAS and CMS collaborations, Combined measurement of the Higgs Boson mass in pp collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS experiments. Phys. Rev. Lett. **114**, 191–803 (2015)
7. L.D. Landau, On the angular momentum of a two-photon system, Dokl. Akad. Nauk Ser. Fiz. **60**, 207 (1948); C.-N. Yang, Selection rules for the dematerialization of a particle into two photons. Phys. Rev. **77**, 242 (1950)
8. ATLAS collaboration, Study of the spin and parity of the Higgs boson in Diboson decays with the ATLAS detector. Eur. Phys. J. C. **75**, 476 (2015)
9. ATLAS collaboration, Combined measurements of Higgs boson production and decay in the $H \rightarrow ZZ \rightarrow 4\ell$ and $H \rightarrow \gamma\gamma$ channels using $\sqrt{s} = 13$ TeV proton-proton collision data collected with the ATLAS experiment, ATLAS-CONF-2017-047
10. ATLAS and CMS collaborations, Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV. JHEP **08**, 045 (2016)
11. ATLAS collaboration, Evidence for the $H \rightarrow b\bar{b}$ decay with the ATLAS detector. [arXiv:1708.03299](https://arxiv.org/abs/1708.03299)

12. ATLAS collaboration, Search for the associated production of a Higgs boson and a top quark pair in multilepton final states with the ATLAS detector, ATLAS-CONF-2016-058
13. ATLAS public results: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/WebHome#Physics_Groups_Combined_Summary; CMS public results: <http://cms-results.web.cern.ch/cms-results/public-results/publications/>

Chapter 9

Spectroscopy of Adsorbates and the Role of Interfacial Interactions



Guido Fratesi , Elena Molteni  and Giovanni Onida 

Abstract We present two test cases showing how electronic and optical spectroscopy techniques and their theoretical understanding by first principles can provide information on the strength of the adsorbate-substrate electronic interaction at interfaces. Results of electronic band structure calculations for two-dimensional silicon sheets (“silicene”) grown on silver are shown, pointing out that the interaction with the metal substrate is strong enough as to disrupt many of the exceptional properties computed for the free-standing material, or to mitigate the differences between various two-dimensional allotropes. The optical properties of silicon surface functionalized by nucleobase molecules are then discussed. Calculations predict that chemical sensitivity of spectra to molecular modifications can be overwhelmed by variations in the substrate response as induced by the adsorbates.

9.1 Introduction

Surfaces of solids forming interfaces with nanometer-sized adsorbed layers can be used for the growth and/or the study of the adsorbed species [6, 33]. For example, large area samples of graphene can be grown on copper and nickel surfaces [15]; organic molecules can be immobilized at a surface for further characterization [1], which is especially effective if ordered samples are obtained [9, 10, 18]. The strength of the adsorbate-substrate interaction is crucial since a compromise is necessary between a weak and ineffective one and one where the adsorbate properties are spoiled by too strong electronic coupling at the interface. Electronic and optical spectroscopy techniques, and their theoretical understanding at the nanometer length

G. Fratesi (✉) · E. Molteni · G. Onida
Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano,
Via Celoria 16, 20133 Milan, Italy
e-mail: guido.fratesi@unimi.it
URL: <https://sites.google.com/site/guidofratesi/>

scale [3, 30], are especially useful to address these issues, as we exemplify by two test cases from our recent work.

Two-dimensional silicon sheets (“silicene”) can be grown on the surface of silver in various forms, which differ for out-of-plane atomic buckling and registry to the substrate but retain a honeycomb structure analogous to graphene [16, 29]. The peculiar electronic structure of the perfect free-standing film is however disrupted by the strong hybridization between Si and Ag states [38]. By combining our first-principles calculations with angle-resolved photoemission experiments we show that all allotropes display similar electronic bands despite the structural differences, missing the massless Dirac fermions [22, 38]. Optical spectra present the fingerprint of silicene-induced transitions, although now with major participation by silver states and photoinduced charge carriers dynamics consequently approaching typical metal timescales [4].

We investigated silicon surfaces covered by uracil-like nucleobases [36] by simulating the reflectance anisotropy spectra (RAS), that can be used to monitor non-destructively the interface. A characteristic RAS lineshape weakly dependent on the adsorbed species provides the mark of uracil-like adsorption. Differences between nucleobases for the molecular transitions in the visible range are however overwhelmed by modifications in the substrate response.

9.2 Spectra of Silicene Allotropes

The search for novel two-dimensional (2D) materials with exceptional properties goes far beyond the popular case of graphene. While theoretical analyses for many real, potential, or hypothetical candidates are routinely performed with *ab initio* methods, experiments and applications often require a substrate where these are grown. Eventually, the substrate may either be part of the experimental setup or device, or be removed for the applications [40]. Among these materials, the case of two-dimensional Si, named silicene, appears especially intriguing because of the potential ease of integration with standard Si-based electronics. With its stability in *vacuo* first identified theoretically [39], it was experimentally realized by evaporating Si on a Ag(111) surface [17]. A main question was then, if the outstanding properties predicted for the free-standing material and in particular the presence of Dirac cones at the Fermi level, were preserved by the interaction with the substrate. That was subject of debate in the literature, with experimental claims [41] then confuted by theoretical [43] and experimental [22] analyses, to mention a few among the abundant related literature. Silicene was also termed as the “swiss-knife” of 2D electronics because various different allotropes are observed, according to growth conditions, which could in principle exhibit tunable properties. Yet the properties of this elusive material on Ag as well as other substrates are not easily understood and require the concurrent contribution by several research groups and techniques, with results not often unequivocally assessed.

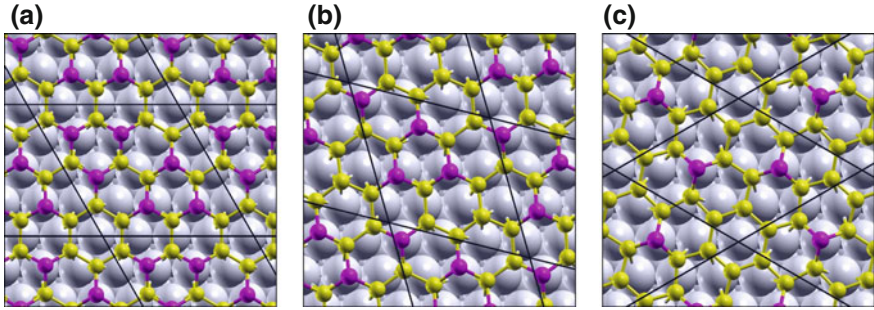


Fig. 9.1 Structural configurations of silicene allotropes on Ag(111): **a** (4×4), **b** ($\sqrt{13} \times \sqrt{13}$), and **c** ($2\sqrt{3} \times 2\sqrt{3}$). Smaller/larger spheres are used for Si/Ag atoms. Protruding Si atoms are shown in magenta (darker)

We applied numerical simulations based on Density Functional Theory (DFT) to the study of various phases of silicene on Ag(111). The slab model was used to describe Ag(111), with Si atoms adsorbed on one side of a five-layer Ag substrate. The calculation cell is periodically repeated also in the direction perpendicular to the substrate and a plane wave basis set with atomic pseudopotentials are used as implemented in the Quantum-ESPRESSO software [11, 12]. We have taken the local density approximation (LDA) to the exchange and correlation functional [32]. Additional details can be found in our earlier works [4, 38].

The studied phases are shown in the ball-stick models of Fig. 9.1a–c referring respectively to a (4×4), ($\sqrt{13} \times \sqrt{13}$), and a ($2\sqrt{3} \times 2\sqrt{3}$) configuration, taking as a reference the underlying Ag substrate (i.e., the (4×4) case is a silicene-(3×3)/Ag(111)-(4×4), and the other two derive from silicene-($\sqrt{7} \times \sqrt{7}$)). Such configurations differ in the relative orientation of the silicene and Ag lattices, and as a consequence in the number and position of Si atoms being atop a surface Ag atom (shown in magenta in Fig. 9.1), with a different buckling and periodicity of the 2D layer. They can be distinguished by scanning tunneling microscopy, Raman spectroscopy [5], and low-energy electron diffraction [38]. Ideally one would like to exploit such differences in device fabrications.

To see how the electronic structure depends on the actual configuration, we start by considering the electronic density. This is compared in Fig. 9.2a, d for the (4×4) and the ($2\sqrt{3} \times 2\sqrt{3}$) structure. Regardless of the configuration taken, one can visualize the accumulation of electronic density along the Si-Si bonds that is also typical of bulk Si, and a metallic behavior underneath. More insight could be gained by analyzing instead the Si-Ag bond charge, defined as

$$\Delta q(\mathbf{r}) = \varrho_{\text{Si/Ag}}(\mathbf{r}) - \varrho_{\text{Si}}(\mathbf{r}) - \varrho_{\text{Ag}}(\mathbf{r}), \quad (9.1)$$

where $\varrho_{\text{Si/Ag}}(\mathbf{r})$, $\varrho_{\text{Si}}(\mathbf{r})$, and $\varrho_{\text{Ag}}(\mathbf{r})$ refer to the electron density of the combined system, the silicene overlayer, and the Ag substrate, respectively (the latter two computed for the same atomic coordinates as in the Si/Ag case). This is shown in

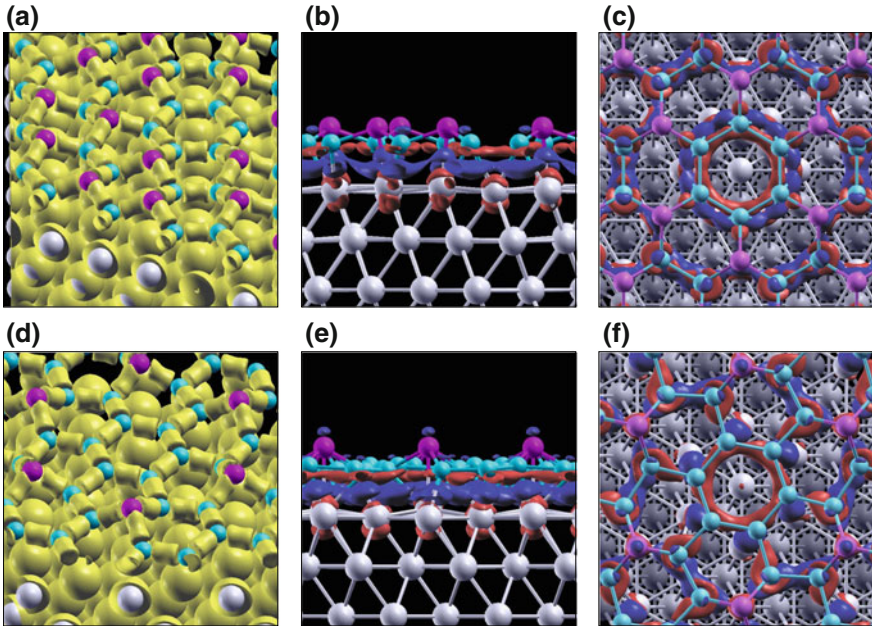


Fig. 9.2 Valence electron density for **a** the (4×4) and **d** $(2\sqrt{3} \times 2\sqrt{3})$ structure. The plotted surfaces correspond to a density of $0.4 e/\text{\AA}^3$. Si-Ag bonding charge, side and top views, for **b**, **c** the (4×4) and **e**, **f** $(2\sqrt{3} \times 2\sqrt{3})$ structure. An isovalue of $\pm 0.04 e/\text{\AA}^3$ is taken, with blue/red regions marking negative/positive charge densities (more/less electrons)

Fig. 9.2b, c, e, f for the same two structures. The overall behavior appears to be similar, with electronic density reduced around the surface Ag atoms and along the Si-Si bonds (see the red regions), and increased in the region between Si and Ag (blue regions). By Löwdin population analysis [21] we estimate a small overall electron displacement towards silicon that becomes more negative by $0.013 e/\text{Si}$ atom and $0.019 e/\text{Si}$ atom in the (4×4) and $(2\sqrt{3} \times 2\sqrt{3})$ case, respectively; such a small difference is not reliable also given the arbitrariness of a charge partitioning scheme.

Even more fundamental is the analysis of the electronic band structure, recalling that the presence of Dirac cones at the K-point of the 2D Brillouin zone (BZ) of free standing silicene (FSS) has been predicted [2, 39]. For optimal tunability and actual device implementation, the possibility to keep the linear dispersion while opening a band gap attracts particular interest, and the different buckling and symmetry of the silicene structures was investigated in this respect. If we keep the same atomic coordinates as upon adsorption, unsupported phases (Ag stripped off) do indeed show remarkable differences as can be seen in Fig. 9.3a–c, where a band gap is visible for the $(4 \times 4)/\text{Ag}$ case in panel (a) at variance with the linear dispersion of the other two cases [13, 34]. Such results can be explained as follows. We recall first that a silicene- (3×3) superlattice [corresponding to the $(4 \times 4)/\text{Ag}$ case in panel (a)] has a 2D BZ that is 9 times smaller than the pristine one with the K point of silicene- (1×1) folded

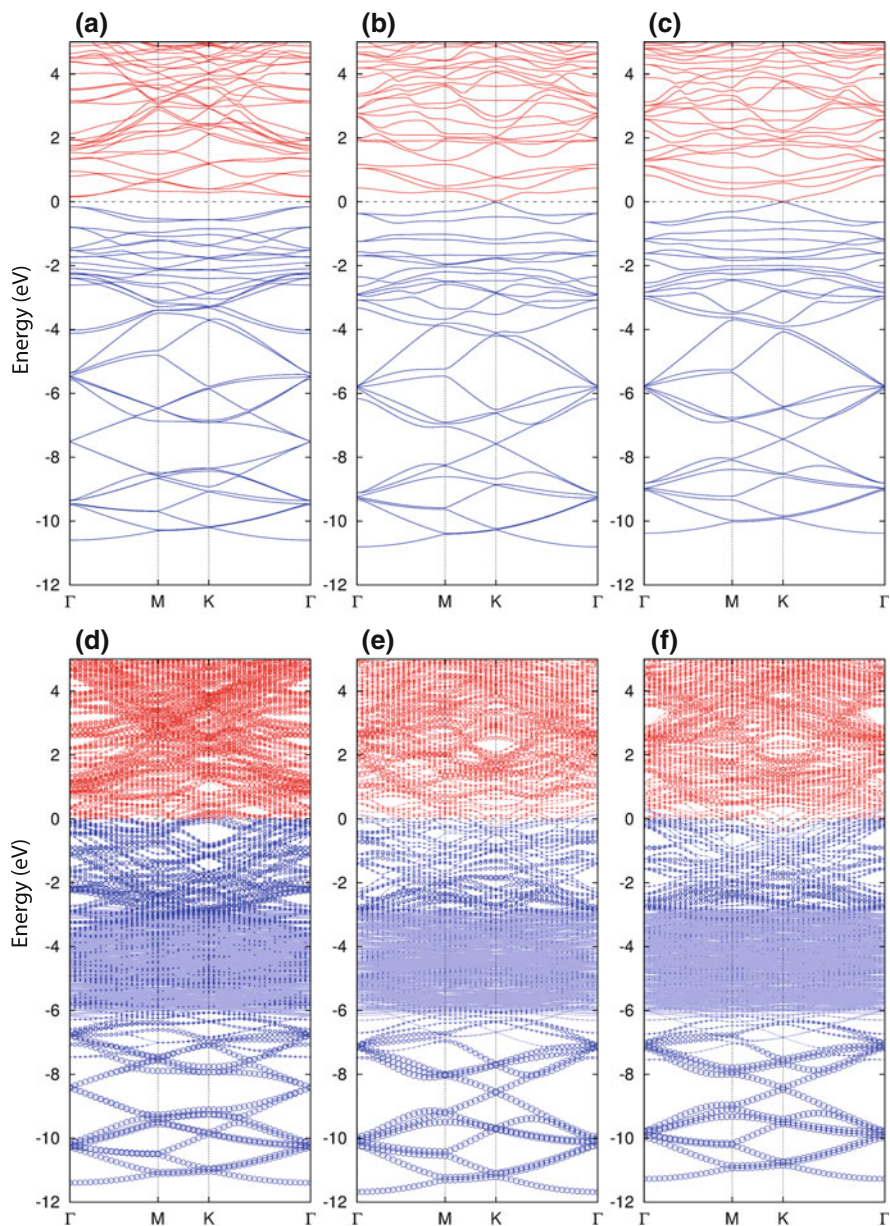


Fig. 9.3 Top row, electronic band structure for unsupported silicene layers (Ag stripped off) arranged with the same atomic coordinates as in the adsorbed case, for **a** (4×4) -unsupported, **b** $(\sqrt{13} \times \sqrt{13})$ -unsupported, and **c** $(2\sqrt{3} \times 2\sqrt{3})$ -unsupported. Blue and red lines are used for filled and empty states, respectively, taking the Fermi energy as a reference. Notice the band gap appearing at the Γ point in **(a)**, and the Dirac cones at K in **(b)** and **(c)**. Bottom row, corresponding band structure computed for the adsorbed cases: **d** $(4 \times 4)/\text{Ag}$, **e** $(\sqrt{13} \times \sqrt{13})/\text{Ag}$, and **f** $(2\sqrt{3} \times 2\sqrt{3})/\text{Ag}$. The circles mark the weight of the states on the Si layer

onto Γ , whereas in the other two cases the K point folds onto the corresponding one of the smaller BZ. On top of this, the electronic potential is perturbed by the modified buckling, that is effective in opening a band gap at the Γ point of the silicene- (3×3) superlattice but not at the K point of the silicene- $(\sqrt{7} \times \sqrt{7})$ one.

To what extent such and other differences are retained in the adsorbed cases was another of the various debated issues concerning this controversial material. To address this issue, we recall facts in the recent literature and present the band structure we compute for the combined system in Fig. 9.3d–f. There, the weight of electronic states on silicene as computed by projecting them on Si-centered orbitals is shown by circles of different sizes. The (4×4) case was the first one where the presence of Dirac cones was ruled out. One can see that the interaction between Si and Ag and the hybridization of their electronic wave functions is so strong as to lose the localization of Si-induced states in favor of hybrid states that bear no specific signature in spectroscopy [22]. Indeed, see the electronic band structure presented in Fig. 9.3d. Clearly, below the Ag d -band (energies lower than about -6 eV), silicene states are mostly unchanged by adsorption. At higher energies they can be hardly followed in the Ag “spaghetti”, as well as at energies above the d -band edge where interaction with the Ag sp -band is remarkable. Later studies focusing on the $(2\sqrt{3} \times 2\sqrt{3})$ case could not establish the presence or absence of the Dirac cone, and proposed a significant shift of the σ bands towards the Fermi energy [42]. As a matter of fact, hardly any difference can be identified by comparing this case, as well as the $(\sqrt{13} \times \sqrt{13})$ one, to the (4×4) in Fig. 9.3. A systematic analysis by angle-resolved photoemission spectroscopy performed consistently for the various phases [38] has recently shown that in all these three cases no state with linear dispersion about the Fermi energy is derived from silicene. Spectral features previously observed could be attributed to Ag or interface states and their umklapp replicas, with differences between the three phases that originate in the relative matching of the Si and Ag superlattices. Remarkably, the quest for exotic features is not yet closed as the presence of six double-cone structures slightly away from the K point has been recently measured [7] and computed [19] for the (4×4) case.

The coupling between metallic and semiconducting systems could be very interesting in devices exploiting, for example, the coupling between plasmonic and excitonic excitations [44]. It is then also important to understand the optical properties of the 2D material and how these are influenced by the interaction with the substrate. To address this issue we present the optical absorbance $A(\omega)$ in Fig. 9.4. That is computed as the light absorption by the 2D material at normal incidence, and within a repeated cell setup it can be derived as [25]:

$$A_x(\omega) = \frac{\omega L_{\perp}}{c} \text{Im}\varepsilon_{xx}(\omega), \quad (9.2)$$

where L_{\perp} is the size of the fictitious cell in the direction perpendicular to the 2D material, and ε_{xx} its dielectric constant with electric field in one of the parallel directions (x). The same quantity can be extracted for free standing and adsorbed cases within the slab model, including in the latter the light absorption by the underlying slab. We

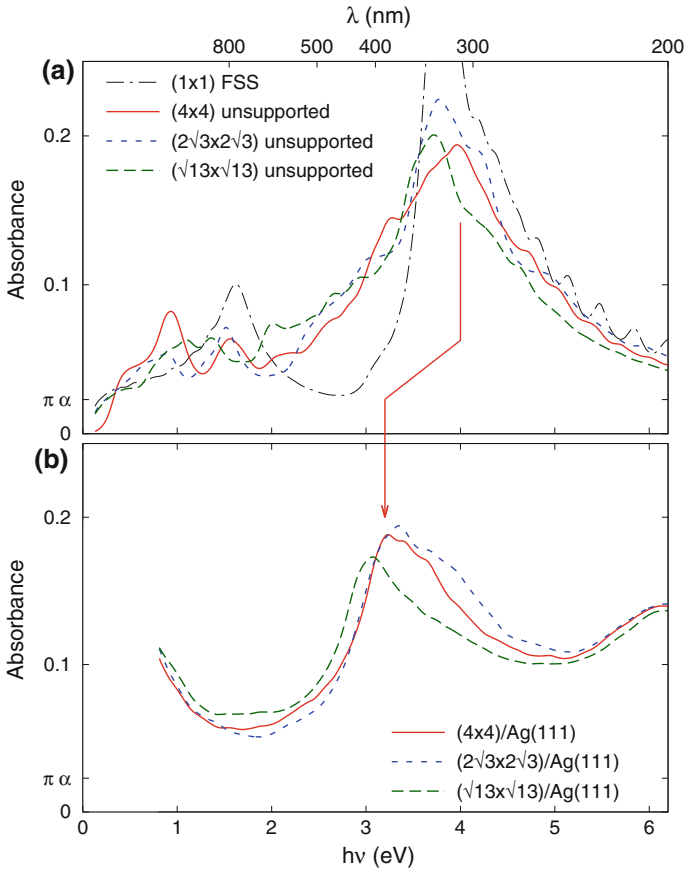


Fig. 9.4 Optical absorbance, see (9.2), for: **a** unsupported silicene phases also compared to the ideal (1×1) FSS; **b** the same systems adsorbed on Ag(111). The red arrow marks the shift in the spectral features of the (4×4) case due to the substrate

compute ε in the independent particle random phase approximation (IP-RPA) using the Yambo software [23].

Our results are presented in Fig. 9.4. We start by the free-standing layers of panel (a). The case of (1×1) silicene, thoroughly studied in the Bechstedt group [25], presents interesting characteristics that are in common with other 2D materials such as graphene and germanene. In particular, one has a universal infrared limit, with $A(\omega \rightarrow 0) = \pi\alpha$ (α being the fine structure constant), and maxima in the absorption that derive from van Hove singularities in the joint density of states (here at 1.6 eV and about 4 eV). This is now compared to the unsupported silicene lattice with the same coordinates as on Ag(111) (see Fig. 9.1). Starting by the (4×4) one, where we recall the results of our previous analysis [4], we can see that the main absorption peaks can still be distinguished, although they appear much less resolved. The zero-

frequency limit is however lost, with $A(\omega \rightarrow 0) = 0$ according to the opening of a band gap as seen in Fig. 9.3a and discussed above. Here we add the results for the other allotropes, $(\sqrt{13} \times \sqrt{13})$ and $(2\sqrt{3} \times 2\sqrt{3})$, which show rather similar overall features but retain the finite zero-frequency limit following the linear dispersion around the Fermi energy.

Upon adsorption, additional modifications are observed in Fig. 9.4b. The spectra broaden further and, while a main silicene-induced peak is still found, it is significantly red-shifted with respect to the unsupported case. Remarkably, no signature of the 1.6 eV peak can be found in the (4×4) phase [4] nor in the other two. At low energy, the spectra are overwhelmed by the Drude absorption peak (not properly included in the current treatment that neglects intraband transitions and not shown here). These results show that the strong hybridization of silicene and Ag states also affects remarkably the optical properties of the 2D layer. Especially, the Si-induced states gain a metallic nature that also reflects in characteristic times of the photogenerated charge carriers more similar to those of bulk Ag than of bulk Si [4]. Possible differences among the phases could be observed, although being smaller than those measurable by Raman spectroscopy [5].

9.3 Optical Anisotropy of Adsorbed Nucleobases

Organic molecule functionalization of Si surfaces could allow the synergistic combination of the flexible properties of organic chemistry with the semiconductor technology. Additionally, binding organic species of biologically relevant function (see DNA/RNA bases) to a surface could ease their analysis, for example for DNA sequencing [8]. For all these applications, the interaction between adsorbed species and the substrate is essential in stabilizing the combined system and obtaining the desired properties. Non-destructive tools of analysis are extremely helpful, and optical spectroscopy would represent a very valuable source of information.

Hence, we now move to a case of organic adsorbates on a semiconductor surface, and address the spectral properties of uracil-like pyrimidinic nucleobases adsorbed on the surface of Si(001). Namely, we consider uracil (URA), thymine (THY), and 5-fluorouracil (5-FU) which differ for the group bound to C at position 5 (see Fig. 9.5). At submonolayer coverage, uracil molecules form ordered structures which have been the subject of experimental [20] and theoretical [28, 37] investigations. A most stable adsorption structure predicts upright molecules adsorbed in the so-called “dimer bridge” configuration [37] as seen in Fig. 9.5 for THY. Rows of facing molecules form, orthogonally to the Si(001) dimers, with alternating (left/right) orientation of the molecular ring.

A natural question is if a spectral analysis can be used to identify the presence of the molecules on the surface, and to distinguish the specific chemical group bound at position 5. The electronic band structure of URA, THY, and 5-FU on Si(001) has been the subject of a previous study by some of us [28]. It was found that the presence of the molecules induces states in the Si band gap that originate from the

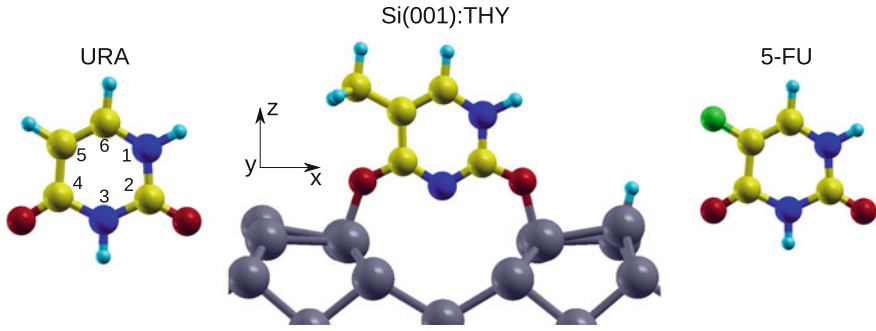


Fig. 9.5 Ball-stick model of uracil-like pyrimidinic nucleobases: uracil (URA), thymine (THY), and 5-fluorouracil (5-FU). For THY, the “dimer bridge” adsorption configuration on Si(001) is shown, with the H atom formerly bound to N at position 3 dissociated and bound to surface Si

molecular LUMO’s in the unoccupied manifold, and from Si states in the occupied one. Remarkably, the electronic band structure in the region of the band gap has shown striking similarities among the three cases. At energies above the conduction band, or below the valence one, differences are not easily identified due to hybridization among molecular and substrate states (recall the discussion of Fig. 9.3).

A valuable non-destructive tool for analysis of surfaces is reflection anisotropy spectroscopy (RAS). The RAS spectra for URA/Si(001) have been first computed by Seino and Schmidt [35] and extended by us to THY and 5-FU [26]. A main effect, common to all three cases, is the flattening of the characteristic “two-horn” RAS of clean Si(001) below ≈ 3 eV. Chemical substitutions in the molecule differentiate the spectra for energies in the near-UV range ($h\nu > 4.5$ eV), at variance with the case of the electronic band structure, whereas they are ineffective in the visible one [26]. Remarkably for the current discussion, such different behavior for high/low photon energy is not inherent in the molecules but results from their interaction with the Si substrate.

We have performed the calculation of the IP-RPA optical properties with the Yambo software [23] on top of Kohn-Sham states computed by Quantum-ESPRESSO [12] with the Perdew-Burke-Ernzerhof (PBE) [31] approximation to the exchange and correlation functional [26, 27]. The RAS was computed by the following expression:

$$\frac{\Delta R}{R}(\omega) = \frac{4\omega L_{\perp}}{c} \text{Im} \left\{ \frac{\varepsilon_{xx}(\omega) - \varepsilon_{yy}(\omega)}{\varepsilon_{\text{bulk}}(\omega) - 1} \right\}, \quad (9.3)$$

which shows clear connection to the expression for the slab absorbance given in (9.2) apart for the complex denominator containing the dielectric function of bulk Si, $\varepsilon_{\text{bulk}}(\omega)$.

Given the similarity between the expressions of $\Delta R/R$ and A , as a simplifying procedure, the absorbance and in particular the absorbance anisotropy $\Delta A = A_x - A_y$ could still be used as a tool with a more straightforward interpretation than the

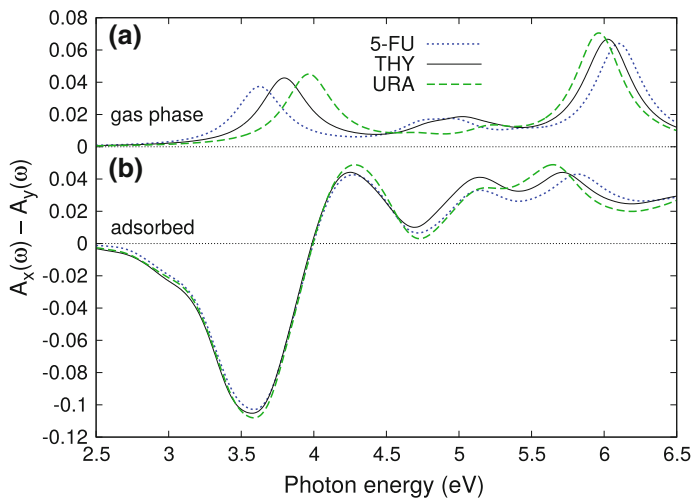


Fig. 9.6 Absorbance anisotropy **a** for gas phase molecules and **b** for nucleobases adsorbed on Si(001)

full RAS, keeping in mind that the complex denominator could change the sign in the RAS features as well as the relative weight of contributions coming from $\text{Re}\varepsilon$ and $\text{Im}\varepsilon$ [14, 26]. While it would bear by itself no direct physical meaning, the same quantity ΔA can be formally defined for the free molecules as a reference (we remark that in this case ΔA is directly proportional to the difference of the photoabsorption cross section in the x-y directions).

We report ΔA for gas phase nucleobases in Fig. 9.6a. One can see that the spectra present a noticeable chemical sensitivity, especially in the 3.5–4 eV range where the transition between the highest occupied and lowest unoccupied molecular orbitals (HOMO-LUMO) is computed: indeed, the energy of the HOMO increases in passing from URA to THY to 5-FU, hence decreasing the excitation energy, while the LUMO stays approximately constant [27]. A milder sensitivity is observed at higher energies, and especially for the peak at about 6 eV which shows the opposite trend.

Moving now to the adsorbed species, the results shown in Fig. 9.6b differ substantially. In particular, in the region of the HOMO-LUMO transition a peak with opposite sign is computed. Remarkably, neither its position nor its magnitude are affected by the specific molecule considered. Detailed analysis of the spatial localization of the electronic states contributing to this peak allowed us to identify it as originating purely from the substrate as a consequence of adsorption [26]. Remarkably, a similar spectral feature could be found by removing the molecules and saturating the cleaved bonds by H atoms. Only at higher energy do modifications in the molecules translate one-to-one in changes of the spectra of the adsorbed system, see the ≈ 6 eV peak (yet with a red-shift).

9.4 Conclusions

Experimental and theoretical approaches to the spectral properties, electronic and optical, can be combined to provide valuable insight into the electronic interaction and hybridization at adsorbate-surface interfaces. We have exemplified this by considering a fully inorganic interface between a novel 2D material (silicene) and a metal surface (Ag), and a hybrid one with nucleobase molecules adsorbed on Si(001). The role of the interface in this respect is generally two-fold: for example, the substrate modification to adsorbate states was detailed for silicene/Ag(111), whereas the adsorbate-induced modifications to the substrate response were shown for the molecules on Si(001).

Especially when the predictive character of first-principle simulations is used, the basic mechanisms and processes occurring at the interfaces can be worked out, paving the ground for optimization of the properties at this crucial ingredient in many systems and devices.

Acknowledgements E.M. acknowledges financial support from Regione Autonoma della Sardegna under Project R.A.S. L.R. 07/08/2007 CRP-26666, and from Fondazione di Sardegna under Project UniCa “PRID 2015”. The authors acknowledge the CINECA award under the ISCRa initiative, for the availability of high-performance computing resources and support (projects HP10CDF3LP and HP10CESYLM). G.O. acknowledges ETSF-Italy for computational resources [24].

References

1. A. Brambilla, A. Picone, D. Giannotti, A. Calloni, G. Berti, G. Bussetti, S. Achilli, G. Fratesi, M. Trioni, G. Vinai, P. Torelli, G. Panaccione, L. Duò, M. Finazzi, F. Ciccacci, Enhanced magnetic hybridization of a spinterface through insertion of a two-dimensional magnetic oxide layer. *Nano Lett.* **17**(12), 7440–7446 (2017). <https://doi.org/10.1021/acs.nanolett.7b03314>
2. S. Cahangirov, M. Topsakal, E. Aktürk, H. Sahin, S. Ciraci, Two- and one-dimensional honeycomb structures of silicon and germanium. *Phys. Rev. Lett.* **102**(23), 236804 (2009). <https://doi.org/10.1103/PhysRevLett.102.236804>
3. A. Calloni, G. Fratesi, S. Achilli, G. Berti, G. Bussetti, A. Picone, A. Brambilla, P. Folegati, F. Ciccacci, L. Duò, Combined spectroscopic and ab initio investigation of monolayer-range Cr oxides on Fe(001): the effect of ordered vacancy superstructure. *Phys. Rev. B* **96**(8), 085427 (2017). <https://doi.org/10.1103/PhysRevB.96.085427>
4. E. Cinquanta, G. Fratesi, S. Dal Conte, C. Grazianetti, F. Scotognella, S. Stagira, C. Vozzi, G. Onida, A. Molle, Optical response and ultrafast carrier dynamics of the silicene-silver interface. *Phys. Rev. B* **92**(16), 165427 (2015). <https://doi.org/10.1103/PhysRevB.92.165427>
5. E. Cinquanta, E. Scalise, D. Chiappe, C. Grazianetti, B. Van Den Broek, M. Houssa, M. Fanciulli, A. Molle, Getting through the nature of silicene: an sp²-sp³ two-dimensional silicon nanosheet. *J. Phys. Chem. C* **117**(32), 16719–16724 (2013). <https://doi.org/10.1021/jp405642g>
6. E. Del Castillo, F. Cargnoni, S. Achilli, G. Tantardini, M. Trioni, Spin asymmetric band gap opening in graphene by fe adsorption. *Surf. Sci.* **634**, 62–67 (2015). <https://doi.org/10.1016/j.susc.2014.11.012>
7. Y. Feng, D. Liu, B. Feng, X. Liu, L. Zhao, Z. Xie, Y. Liu, A. Liang, C. Hu, Y. Hu, S. He, G. Liu, J. Zhang, C. Chen, Z. Xu, L. Chen, K. Wu, Y.T. Liu, H. Lin, Z.Q. Huang, C.H. Hsu,

- F.C. Chuang, A. Bansil, X.J. Zhou, Direct evidence of interaction-induced Dirac cones in a monolayer silicene/Ag(111) system. *Proc. Natl. Acad. Sci. U.S.A.* **113**(51), 14656–14661 (2016). <https://doi.org/10.1073/pnas.1613434114>
8. D. Fologea, M. Gershow, B. Ledden, D.S. McNabb, J.A. Golovchenko, J. Li, Detecting single stranded DNA with a solid state nanopore. *Nano Lett.* **5**(10), 1905–1909 (2005). <https://doi.org/10.1021/nl051199m>
 9. G. Fratesi, V. Lanzilotto, L. Floreano, G.P. Brivio, Azimuthal dichroism in near-edge x-ray absorption fine structure spectra of planar molecules. *J. Phys. Chem. C* **117**(13), 6632–6638 (2013). <https://doi.org/10.1021/jp312569q>
 10. G. Fratesi, V. Lanzilotto, S. Stranges, M. Alagia, G.P. Brivio, L. Floreano, High resolution NEXAFS of perylene and PTCDI: a surface science approach to molecular orbital analysis. *Phys. Chem. Chem. Phys.* **16**(28), 14834–14844 (2014). <https://doi.org/10.1039/c4cp01625d>
 11. P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M.B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. Dal Corso, S. de Gironcoli, P. Delugas, R.A. DiStasio Jr., A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H.Y. Ko, A. Kokalj, E. Küçükbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N.L. Nguyen, H.V. Nguyen, A. Otero-de-la Roza, L. Paulatto, S. Poncé, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A.P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu, S. Baroni, Advanced capabilities for materials modelling with Quantum ESPRESSO. *J. Phys. Condens. Matter* **29**(46), 465901 (2017). <https://doi.org/10.1088/1361-648X/aa8f79>
 12. P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G.L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. De Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A.P. Seitsonen, A. Smogunov, P. Umari, R.M. Wentzcovitch, Quantum ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys. Condens. Matter* **21**(39), 395502 (2009). <https://doi.org/10.1088/0953-8984/21/39/395502>
 13. Z.X. Guo, S. Furuya, J.I. Iwata, A. Oshiyama, Absence and presence of Dirac electrons in silicene on substrates. *Phys. Rev. B* **87**(23), 235435 (2013). <https://doi.org/10.1103/PhysRevB.87.235435>
 14. C. Hogan, R. Del Sole, Optical properties of the GaAs(001)-c(4 × 4) surface: direct analysis of the surface dielectric function. *Phys. Status Solidi B Basic Res.* **242**(15), 3040–3046 (2005). <https://doi.org/10.1002/pssb.200562231>
 15. J. Jang, M. Son, S. Chung, K. Kim, C. Cho, B.H. Lee, M.H. Ham, Low-temperature-grown continuous graphene films from benzene by chemical vapor deposition at ambient pressure. *Sci. Rep.* **5** (2015). <https://doi.org/10.1038/srep17955>
 16. A. Kara, H. Enriquez, A.P. Seitsonen, L.C. Lew Yan Voon, S. Vizzini, B. Aufray, H. Oughaddou, A review on silicene—new candidate for electronics. *Surf. Sci. Rep.* **67**(1), 1–18 (2012). <https://doi.org/10.1016/j.surfrep.2011.10.001>
 17. B. Lalmi, H. Oughaddou, H. Enriquez, A. Kara, S. Vizzini, B. Ealet, B. Aufray, Epitaxial growth of a silicene sheet. *Appl. Phys. Lett.* **97**(22), 223109 (2010). <https://doi.org/10.1063/1.3524215>
 18. V. Lanzilotto, C. Sanchez-Sanchez, G. Bavdek, D. Cvetko, M.F. Lopez, J.A. Martin-Gago, L. Floreano, Planar growth of pentacene on the dielectric TiO₂(110) surface. *J. Phys. Chem. C* **115**(11), 4664–4672 (2011). <https://doi.org/10.1021/jp111011z>
 19. C. Lian, S. Meng, Dirac cone pairs in silicene induced by interface Si-Ag hybridization: a first-principles effective band study. *Phys. Rev. B* **95**(24), 245409 (2017). <https://doi.org/10.1103/PhysRevB.95.245409>
 20. A. Lopez, Q. Chen, N.V. Richardson, Combined STM, HREELS and ab initio study of the adsorption of uracil on Si(100)-2 × 1. *Surf. Interface Anal.* **33**(5), 441–446 (2002). <https://doi.org/10.1002/sia.1231>

21. P.O. Löwdin, On the non-orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals. *J. Chem. Phys.* **18**(3), 365 (1950). <https://doi.org/10.1063/1.1747632>
22. S.K. Mahatha, P. Moras, V. Bellini, P.M. Sheverdyeva, C. Struzzi, L. Petaccia, C. Carbone, Silicene on Ag(111): a honeycomb lattice without Dirac bands. *Phys. Rev. B* **89**(20), 201416 (2014). <https://doi.org/10.1103/PhysRevB.89.201416>
23. A. Marini, C. Hogan, M. Grüning, D. Varsano, yambo: an ab initio tool for excited state calculations. *Comput. Phys. Commun.* **180**(8), 1392–1403 (2009). <https://doi.org/10.1016/j.cpc.2009.02.003>
24. A. Matsuura, N. Thrupp, X. Gonze, Y. Pouillon, G. Bruant, G. Onida, The ETSF: an e-infrastructure that bridges simulations and experiments. *Comput. Sci. Eng.* **14**(1), 22–32 (2012). <https://doi.org/10.1109/MCSE.2011.76>
25. L. Matthes, P. Gori, O. Pulci, F. Bechstedt, Universal infrared absorbance of two-dimensional honeycomb group-IV crystals. *Phys. Rev. B* **87**(3), 035438 (2013). <https://doi.org/10.1103/PhysRevB.87.035438>
26. E. Molteni, G. Cappellini, G. Onida, G. Fratesi, Optical properties of organically functionalized silicon surfaces: uracil-like nucleobases on Si(001). *Phys. Rev. B* **95**(7), 075437 (2017). <https://doi.org/10.1103/PhysRevB.95.075437>
27. E. Molteni, G. Fratesi, G. Cappellini, G. Onida, Optical properties of free and Si(001)-adsorbed Pyrimidinic Nucleobases. *Phys. Status Solidi B*, 1700497 (2017). <https://doi.org/10.1002/pssb.201700497>
28. E. Molteni, G. Onida, G. Cappellini, Electronic structure of uracil-like nucleobases adsorbed on Si(001): uracil, thymine and 5-fluorouracil. *Eur. Phys. J. B* **89**(4), 98 (2016). <https://doi.org/10.1140/epjb/e2016-70011-1>
29. H. Oughaddou, H. Enriquez, M.R. Tchalala, H. Yildirim, A.J. Mayne, A. Bendounan, G. Dujardin, M. Ait Ali, A. Kara, Silicene, a promising new 2D material. *Prog. Surf. Sci.* **90**(1), 46–83 (2015). <https://doi.org/10.1016/j.progsurf.2014.12.003>
30. S. Pagliara, S. Tognolini, L. Bignardi, G. Galimberti, S. Achilli, M. Trioni, W. Van Dorp, V. Ocelk, P. Rudolf, F. Parmigiani, Nature of the surface states at the single-layer graphene/Cu(111) and graphene/polycrystalline-Cu interfaces. *Phys. Rev. B* **91**(19), 195440 (2015). <https://doi.org/10.1103/PhysRevB.91.195440>
31. J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**(18), 3865–3868 (1996). <https://doi.org/10.1103/PhysRevLett.77.3865>
32. J.P. Perdew, A. Zunger, Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B* **23**(10), 5048–5079 (1981). <https://doi.org/10.1103/PhysRevB.23.5048>
33. A. Picone, M. Riva, G. Fratesi, A. Brambilla, G. Bussetti, M. Finazzi, L. Duò, F. Ciccacci, Enhanced atom mobility on the surface of a metastable film. *Phys. Rev. Lett.* **113**(4), 046102 (2014). <https://doi.org/10.1103/PhysRevLett.113.046102>
34. E. Scalise, E. Cinquanta, M. Houssa, B. Van Den Broek, D. Chiappe, C. Grazianetti, G. Pourtois, B. Ealet, A. Molle, M. Fanciulli, V.V. Afanas'ev, A. Stesmans, Vibrational properties of epitaxial silicene layers on (1 1 1) Ag. *Appl. Surf. Sci.* **291**, 113–117 (2014). <https://doi.org/10.1016/j.apsusc.2013.08.113>
35. K. Seino, W.G. Schmidt, Reflectance anisotropy of uracil covered Si(0 0 1) surfaces: ab initio predictions. *Surf. Sci.* **548**(1–3), 183–186 (2004). <https://doi.org/10.1016/j.susc.2003.11.002>
36. K. Seino, W.G. Schmidt, F. Bechstedt, Organic modification of surface electronic properties: a first-principles study of uracil on Si(001). *Phys. Rev. B* **69**(24), 245309 (2004). <https://doi.org/10.1103/PhysRevB.69.245309>
37. K. Seino, W.G. Schmidt, M. Preuss, F. Bechstedt, Uracil adsorbed on Si(001): structure and energetics. *J. Phys. Chem. B* **107**(21), 5031–5035 (2003). <https://doi.org/10.1021/jp0342531>
38. P.M. Sheverdyeva, S.K. Mahatha, P. Moras, L. Petaccia, G. Fratesi, G. Onida, C. Carbone, Electronic states of silicene allotropes on Ag(111). *ACS Nano* **11**(1), 975–982 (2017). <https://doi.org/10.1021/acsnano.6b07593>

39. K. Takeda, K. Shiraishi, Theoretical possibility of stage corrugation in Si and Ge analogs of graphite. *Phys. Rev. B* **50**(20), 14916–14922 (1994). <https://doi.org/10.1103/PhysRevB.50.14916>
40. L. Tao, E. Cinquanta, D. Chiappe, C. Grazianetti, M. Fanciulli, M. Dubey, A. Molle, D. Akinwande, Silicene field-effect transistors operating at room temperature. *Nat. Nanotechnol.* **10**(3), 227–231 (2015). <https://doi.org/10.1038/nnano.2014.325>
41. P. Vogt, P. De Padova, C. Quaresima, J. Avila, E. Frantzeskakis, M.C. Asensio, A. Resta, B. Ealet, G. Le Lay, Silicene: compelling experimental evidence for graphene like two-dimensional silicon. *Phys. Rev. Lett.* **108**(15), 155501 (2012). <https://doi.org/10.1103/PhysRevLett.108.155501>
42. W. Wang, W. Olovsson, R.I.G. Uhrberg, Experimental and theoretical determination of σ bands on $(2\sqrt{3} \times 2\sqrt{3})$ silicene grown on Ag(111). *Phys. Rev. B* **92**(20), 205427 (2015). <https://doi.org/10.1103/PhysRevB.92.205427>
43. Y.P. Wang, H.P. Cheng, Absence of a Dirac cone in silicene on Ag(111): first-principles density functional calculations with a modified effective band structure technique. *Phys. Rev. B* **87**(24), 245430 (2013). <https://doi.org/10.1103/PhysRevB.87.245430>
44. W. Zhang, A.O. Govorov, G.W. Bryant, Semiconductor-metal nanoparticle molecules: Hybrid excitons and the nonlinear Fano effect. *Phys. Rev. Lett.* **97**(14), 146804 (2006). <https://doi.org/10.1103/PhysRevLett.97.146804>

Chapter 10

The Founder of the First Institute of Physics of the Milan University: Aldo Pontremoli, a Physicist's Life Between Adventure and Institutions



Leonardo Gariboldi

Abstract Aldo Pontremoli was the founder of the first Institute of Physics of the University of Milan. The aim of this paper is to offer a reconstruction of his scientific career and how it was affected by the interplay between scientific practices, institutional abilities, and love for adventure. Pontremoli's scientific life will be analysed in three main periods: the training years, the researches in Rome, and the researches in Milan. A particular attention will be given to the establishment of the Institute of Complementary Physics and of the Laboratory of Radiology, to the public competition for the chair of Theoretical Physics, and the 1928 polar expedition on the Italia airship.

10.1 Introduction

Today's Department of Physics of the University of Milan was established after the 1980 reformation of the Italian universities as a transformation of the previous Institute of Physics¹ named after Aldo Pontremoli (1896–1928). A first historiographical interest in studying Pontremoli is therefore a quite local one to be seen, at the same time, as a fundamental step in the reconstruction of the 20th century history of physics in Milan.

Pontremoli was actually the founder of the Institute of Complementary Physics in 1924 following the very establishment of the University of Milan in the same year. The first Rector, the physician Luigi Mangiagalli (1850–1928), who was at

¹On the Institute of Physics, see [2, 16, 59].

L. Gariboldi (✉)

Dipartimento di Fisica "Aldo Pontremoli", Università degli Studi di Milano, Milan, Italy
e-mail: leonardo.gariboldi@unimi.it

the same time Mayor of Milan and Senator of the Kingdom, was a strong supporter of the Faculty of Science, where students could graduate in Mathematics, Natural Sciences, and Industrial Chemistry. An agreement with the University of Pavia but forbade the possibility to graduate students in Physics. A second historiographical topic is thus the reconstruction and analysis of: the role played by Pontremoli in creating a new Institute of Physics from nothing; the structure of the Institute; the scientific and didactical activities carried on there.

During the four last years of his life, Pontremoli was able to favour his scientific interests thanks to the connections his family had with the financial milieu in Milan. His ability in not fully depending from the financial support granted by the University, but also in relying on adequate ones by local private or public patrons, was quite similar, even if on a much higher level, to that of his successor, Giovanni Polvani (1892–1970). If it goes without saying that one's own family affects a scientist's life, it is instead quite a rare fact that it plays such a role in the life of a whole institution. A third topic is the recognition of the "institutional" connections and aspects of Pontremoli's life as active factors in his scientific career.

While in Milan, Pontremoli took part to the public competition for the first Italian chair of Theoretical Physics. It is a rather well known fact that the three winners were Enrico Fermi (1901–1954), Enrico Persico (1900–1969), and Pontremoli, but the latter is always just mentioned as the third and last one. A fourth topic is the analysis of the documents of the public competition in order to better understand why Pontremoli became professor of Theoretical Physics.

Eventually Pontremoli died during the 1928 Italian polar expedition in the Sea of Barents. His role in planning and participating to the first entirely scientific polar expedition on airship cannot be analysed without considering his passion for Alpine exploration and his participation to the First World War as a military officer on balloon. A last topic is seeing how Pontremoli's passion for scientific research, adventure, and exploration tangled up.

10.2 The Training Years

Aldo Pontremoli² was born in Milan on the 19th of January 1896. He was the son of Alfredo Pontremoli (1865–1911) and Lucia Luzzatti (1867–1957), a family of Jewish descent. His father was an engineer, director of a paper mill in Besozzo, a location close to the Maggiore Lake; he was very interested in the sciences and arts.

²Primary sources are held in the following archives: Archivio di Stato in Rome, Historical Archives of Banca Intesa Sanpaolo in Milan, Historical Archives of Banca Popolare di Milano in Milan, Luigi Luzzatti Archive of Istituto Veneto di Scienze Lettere ed Arti in Venice. Short biographies can be found in [12, 14, 15, 51, 52, 54]. Many biographical non-scientific details can be found in [18], a book written after Pontremoli's death and mostly based on the family's memoirs and the written correspondence between Pontremoli and his lifelong friend Massimiliano Majnoni d'Intignano (1894–1957) held in the Historical Archives of Banca Intesa Sanpaolo. Pontremoli's mother thought the way her son's life was described was a faithful one.

His mother was the daughter of Luigi Luzzatti (1841–1927), a person who played an impressive influence on Pontremoli. Luzzatti was professor of Constitutional Law at the University of Padua in 1867–1896, the founder and first president of the Banca Popolare di Milano in 1865, the supporter to the establishment of other popular banks in Italy, one of the founders of the University of Venice, an expert in the history of religions, minister in various governments, and eventually President of the Council of Ministers of the Kingdom in 1910–1911. Pontremoli's family milieu shaped his life from the beginning: a deep interest in science and everyday connections with the industrial, cultural, political and financial world.

Pontremoli soon showed his abilities in merging scientific and institutional practices. He was extremely interested in the contemporary researches on radio-waves transmission. He therefore got in touch with Guglielmo Marconi (1874–1937). As a ten years old child Pontremoli founded with two of his friends a small “radiotelegraphy society” in order to study how a radio wave device worked.³ This society was not just a young teenager's pastime, since it lasted until he began to study physics at the Beccaria Lyceum in Milan. His physics professor was Temistocle Calzecchi Onesti (1879–1922). In 1884 Calzecchi Onesti started to study the conductivity of excited metallic powders. He thus invented a glass tube filled with nickel and silver powders whose conductivity increased when struck by electromagnetic waves. Calzecchi Onesti used this tube to improve the Hughes-Bell telephone and a micro-seismic detector (see [3–5]). His tube was later called a “coherer” by Oliver Lodge (1851–1940) and was improved by Marconi to become a fundamental part of his radio wave apparatus as a detector of electromagnetic waves. Having Calzecchi Onesti as his physics professor moved Pontremoli to the decision to study physics in a proper way, so that he decided to close his radiotelegraphy society.

At that time the only place to study physics at a university level in Milan was the Royal Polytechnic High School. Pontremoli enrolled as a student of Engineering. His passion for scientific practices made him reproduce at home the experiments he took part in at the polytechnic laboratories. His physics professor was Oreste Murani (1853–1937) whose main fields of research were the transmission of electromagnetic waves, the polarisation of light, the ionisation of gases, and the behaviour and properties of X-rays.⁴

On the 24th of May 1915 Italy entered the First World War. Pontremoli stopped his university studies and volunteered in the Army since he was a convinced interventionist. His military life started in a balloon division close to Padua. He was later moved to Udine as the director of the observational tower on top of the castle. On the 30th of July 1915 his practical training on the war front began. He learned how to fly a drachen, the military tethered balloon, and draw geographic surveys of the Austro-

³Teenagers, usually boys, building a radio apparatus by themselves were not an unusual fact until a few decades ago. We can never highlight enough the fact that Pontremoli, although in a childish way, framed in an institutional structure his scientific passion in building devices to study radio-waves transmission. Such an organisational mind was most probably the effect of his family education.

⁴Pontremoli always felt a profound esteem for Murani. When Pontremoli came back to Milan in 1924 as a university professor, he refused to replace Murani and teach physics also at the Polytechnical High School.

Hungarian army stations on the Venete front. He was promoted second lieutenant of the Balloon Military Engineering Corps of the Third Army in the Fifth Balloon Section, later in the Fourth (March 1916) and the Third sections (1917). After Italy's defeat in Caporetto, the Third Balloon Section was shifted back to Ferrara. In 1918 Pontremoli asked and succeeded in being sent to the French front where he fought in the Ardennes and in the Bois de Bligny. At the end of November 1918 he was sent to Belgium for a short time before coming back to Milan. For his military merits he was awarded with a silver medal, a military cross, one French and two Italian war crosses.

In 1919 Pontremoli could start again his university studies. He enrolled at the fourth year of Physics at the University of Rome. In this way he could often spend some time with his grandfather. On the 3rd of July 1921, as a student of Orso Mario Corbino (1876–1937), Pontremoli graduated in Physics with full marks *cum laude*. His dissertation thesis was a deep study of the birefracton fringes in a liquid in a permanent plane motion [27, 28]. Accidental birefracton in moving liquids had been first detected by Laureto Tieri (1879–1952) in 1910. It could be observed in solutions of Bravais colloidal iron. This substance was at disposal in Corbino's laboratory since he used it in his researches on magnetic fields. Pontremoli built an experimental device to study birefracton in the effluent of a colloidal iron solution from capillary tubes. He also formulated a theory of birefracton in viscous liquids, based on Carl Neumann's theory of birefracton in isotropic solids. He deduced the existence of two fringe families. If we consider a monochromatic radiation of wavelength λ striking a liquid layer with thickness s_z , refraction index n (when undeformed), elastic-optical constant p , relaxation time T , in motion at a velocity \mathbf{v} , then its birefracton index in the z direction is given by Pontremoli's formula:

$$\Delta_z = \frac{\pi s_z n^3}{\lambda c^2} p T \sqrt{\left(\frac{\partial v_x}{\partial x} - \frac{\partial v_y}{\partial y}\right)^2 + \left(\frac{\partial v_x}{\partial y} + \frac{\partial v_y}{\partial x}\right)^2} \quad (10.1)$$

The first fringe family depends on the angle Φ_z between the fixed frame of reference and the local dilation principal axes Ox_0 , Oy_0 , given by the relation:

$$\tan^2 \Phi_z = \left(\frac{\partial v_x}{\partial y} + \frac{\partial v_y}{\partial x}\right) / \left(\frac{\partial v_x}{\partial x} - \frac{\partial v_y}{\partial y}\right) \quad (10.2)$$

From the formula connecting the intensity of the incident radiation I with that of the emerging radiation I_0 from a birefracting material:

$$I = I_0 \sin^2 2(\Phi_z - \phi) \sin^2 \frac{\Delta_z}{2} \quad (10.3)$$

where ϕ is the angle between the axes of the fixed frame of reference and the projections on the fixed XY plane of the polariser and analyser principal sections, Pontremoli obtained a system of two equations corresponding to two fringe families.

As noticed above, the first fringe family depends only on the position of the polariser principal section and from the deformation speed of the liquid:

$$\Phi_z - \phi = \frac{k\pi}{2} \quad (10.4)$$

The second fringe family does not depend on the position of the polarisation plane but is a function of the physical properties of the liquid, λ , s_z , and \mathbf{v} :

$$\Delta_z = 2k\pi \quad (10.5)$$

This first research by Pontremoli was on a subject of late classical physics. Apart from taking into account a generic atomic structure of matter, no quantum considerations were made. To be a modern physicist required a different kind of work. If Italian Mathematics was excellent around 1920, the same cannot be said for Italian Physics (see, e.g. [20], p. 175). The number of physicists,⁵ the quality of researches, journals⁶ and conferences, the financial support were all factors which can not be compared to the correspondent ones in the most advanced European countries and in the United States. With the exception of astrophysical spectroscopy (in the astronomical observatories, not in the universities), Italian researches on X-rays, radioactivity, cosmic rays were not abundant enough to compete with the researches carried on in other countries. In order to improve his modern physics education, Pontremoli went to the Cavendish Laboratory in Cambridge to study atomic and nuclear physics thanks to a 8,000 lire scholarship from the Associazione Nazionale Combattenti. Coming from a peripheral country, Pontremoli found himself in one of the beating hearts of contemporary physics research. In Cambridge, Pontremoli was trained in experimental research and radioactivity measurements and attended lectures given by Joseph Larmor (1857–1942), Joseph John Thomson (1856–1940), Arthur Eddington (1882–1944), Ernest Rutherford (1871–1937), and Charles Galton Darwin (1887–1962). He cooperated in the counting operations of atoms recoiling from a homogeneous α ray beam [23]. According to Pugno Vanoni ([52], p. 42) Rutherford wrote words of appreciation of his “great fertility in speculation and suggestions for further experiment.”⁷ Pontremoli was expected to stay in Cambridge for one or two years, but there was no other scholarship to further support his studies in Cambridge.

⁵There was usually just one physicist in each important university. Rome was an exception with two physicists.

⁶The most important Italian physics journal, *Il Nuovo Cimento* became an international journal only after the Second World War.

⁷Rutherford’s document has not been identified yet.

10.3 Pontremoli's Researches in Rome

Pontremoli came back from Cambridge to Rome to work as Corbino's assistant. From the end of the First World War, Corbino was less and less an active physicist; even if he continued to teach and to direct the Institute of Physics of the University of Rome (in such a role, he was the later supporter of the Via Panisperna School, i.e. Fermi's research team), he spent most of his time as a Senator of the Kingdom. Two other physicists started to work with Corbino at that time: Enrico Fermi and Enrico Persico. Just like Pontremoli, both of them would have gone abroad to some of the centres of contemporary physics: Fermi to Göttingen and Leiden in 1922–1923, Persico to Cambridge in 1925. A new generation of young physicists was growing to modernise Italian physics and shifting it from a peripheral condition to a central one in the future.

In Rome, Pontremoli carried on researches on both classical and modern physics. After his work on birefracton, the first subject he studied was how to modify Maxwell's equations of a linearly polarised wave propagating inside a dielectric material when the latter is placed in two longitudinal electric and magnetic fields [29]. A similar study had been done on a conductor (an iron wire) by Gustav Wiedemann put inside two longitudinal electric and magnetic fields. Inspired by Curie's theory of symmetries, Pontremoli suggested the existence of optical phenomena in an isotropic material placed inside electric and magnetic longitudinal fields. He considered a material with N symmetric molecules per unit volume, with one polarisation electron each. These electrons are moved away from the equilibrium position by the action of the external constant electric ($\mathbf{E} = E\mathbf{e}_z$) and magnetic ($\mathbf{H} = H\mathbf{e}_z$) fields and by the action of the variable electric field ($\mathbf{E}_{\text{rad}} = E_x\mathbf{e}_x + E_y\mathbf{e}_y$) of the radiation (linearly polarised in the xy plain). He obtained a formula for the polarisation \mathbf{P} .

$$\mathbf{P} = \frac{\theta N}{4\pi} \frac{\Delta E_x + i\psi E_y}{\Delta^2 - \psi^2} \mathbf{e}_x + \frac{\theta N}{4\pi} \frac{\Delta E_y - i\psi E_x}{\Delta^2 - \psi^2} \mathbf{e}_y + \frac{\theta N}{4\pi} E\mathbf{e}_z \quad (10.6)$$

where r and θ are positive constants, $T_P = 2\pi\tau_P$ is the resonance period of the polarisation electron, $\Delta = 1 + ia/\tau - b/\tau^2$, $a = r\sigma/4\pi$, $b = m\theta/4\pi e^2 = \tau_P^2$, $\psi = \sigma H/4\pi c\tau e$. The motion of the polarisation electrons causes a magnetisation $\mathbf{M} = (\mathbf{P} \times \mathbf{v})/2c$ to be introduced in Maxwell's equations, leading to Pontremoli's equation:

$$\frac{\partial^2 \mathbf{E}_{\text{rad}}}{\partial t^2} + 4\pi \frac{\partial^2 \mathbf{P}}{\partial t^2} + 4\pi c \nabla \times \frac{\partial \mathbf{M}}{\partial t} = c^2 \frac{\partial^2 \mathbf{E}_{\text{rad}}}{\partial z^2} \quad (10.7)$$

Pontremoli's equation describes two circularly polarised waves which propagate parallel to the lines of force of the external fields. The rotating power which modified the radiation polarisation was actually created by the two longitudinal external fields. From this rotating power, Pontremoli could obtain the refraction index and the absorption coefficients for both waves [30]. The action of a longitudinal electric field thus modified both the velocity of propagation of circularly polarised waves

(via a modification of the refraction indexes) and the absorption coefficients, with consequences in spectroscopic analysis.

The first research Pontremoli carried on in modern physics in Rome was a theoretical study on Rutherford's neutron⁸ [17]. Rutherford's 1920 Bakerian Lecture [55] impressed Pontremoli. Rutherford had advanced the existence of a proton-electron compound neutral structure and of possible more complex neutral compact structures inside the atomic nucleus. In Cambridge, Glasson attempted to detect Rutherford's neutrons in hydrogen-filled discharge tubes [19], with a negative result which could not be considered a definitive one, since Rutherford's neutron's properties were completely unknown. Pontremoli analysed Rutherford's neutron in the framework of Bohr's atomic theory [31].

In a first model, Pontremoli criticised the proposal of Rutherford's neutron as a hydrogen atom whose electron was on an orbit extremely close to the proton. He suggested that the proton and the electron needed to be tangent to each other, since Bohr's theory forbade a hydrogen atom smaller than that on the fundamental state. With Silberstein's formula [56] of the mass defect of a system of two spherical electric distributions of charges e_1 and e_2 , radiuses r_1 and r_2 , and whose centres are at a distance d :

$$\Delta m = \frac{e_1 e_2}{d c^2} \left(1 - \frac{1}{5} \frac{r_1^2 + r_2^2}{d^2} \right) \quad (10.8)$$

and the ratio between proton and electron masses from spectroscopic measurements [13] Pontremoli found a value $\Delta m = -2.396 \times 10^{-27}$ g, corresponding to $\Delta E = -2.15 \times 10^{-6}$ erg, which should have been emitted as a 0.009 Å wavelength radiation during the creation of one Rutherford's neutron. If this were the real case, Glasson could have never been able to detect the creation of a Rutherford's neutron.

In a second model, Pontremoli considered the existence of more complex compound neutrons and tried a theoretical study of a plane hydrogen nuclei with nuclei and electrons' positions exchanged. With Sommerfeld's formulae [58], Pontremoli calculated some of the parameters describing this kind of Rutherford's neutron. Since the radius of the nuclear orbits was much smaller than that of the electrons in a normal hydrogen molecule, the Balmer emission spectrum would have been in the X-rays range. Pontremoli did not continue the analysis of this model and, to our knowledge, did not attempt to test it with an experiment.

A short note signed by Fermi and Pontremoli [11] was instead a contribution to relativistic physics with some considerations on the electromagnetic mass and on the inertia of charged particles.

Another subject of research chosen by Pontremoli was spectroscopy: he studied some topics concerning the electric discharges in a rarefied gas [34] and the effect of the magnetic field on them. Robert Wood's experiments [60–63] on hydrogen discharge tubes had shown that: (1) the presence of water vapour caused a high level of dissociation; (2) recombination of hydrogen atoms in molecules could be caused

⁸We call these nuclear structures "Rutherford's neutron"—as Pontremoli himself did in the title of his paper—to distinguish them from the real neutron.

by metal traces or by the closeness of the electrodes. Pontremoli conjectured that the influence of water vapour was due to the increasing dissociation of molecular hydrogen inside the tube (in analogy to ionisation in electrolytic phenomena). He evaluated the effect of the electric field on the dissociation degree and on the discharge in the gas; he also conjectured that the electrically charged particles of the dissociated gas moved at a speed much larger than the thermal one. Consequently, the thermodynamic equilibrium changed and the gas dissociation increased. Pontremoli could thus deduce the function of the electrodes or of metal traces: the partial pressure of the electrons in a metal was larger than that of the electrons produced by the gas ionisation; their concentration was therefore different and they tried to escape the electrodes into the gas and change the thermodynamic equilibrium. The electron flux from the electrodes into the gas should have been responsible for the recombination of molecules and atoms.

By placing a U-shaped Wood tube between the two poles of a Weiss electromagnet (about 7,000 gauss), Pontremoli studied a new effect of the magnetic field on the discharge of rarefied gases [32]. The power supplied by a Corbino-Trabacchi apparatus caused a very high level of dissociation in the gas. Pontremoli's experiment showed that: (1) at low pressures, the magnetic field thinned the light beam and intensified the spectrum; (2) at higher pressures (about 1–1.5 mm Hg), the beam did not get thinner and was whiter, while the intensity of the Balmer spectrum decreased when compared to the intensity of the secondary spectrum. The observed phenomenon was quite similar to that with a higher pressure in the part of tube inside the magnetic field.

Quantum physics was applied by Pontremoli also in the study of thermionic emission [33]. Richardson's formula for the intensity of saturation thermionic current as a function of absolute temperature $I = AT^2 e^{-b/T}$ had been proved by Dushman [10] with thermodynamical considerations. Dushman was also able to write the value A (constant for any element) as a function of universal constants. Pontremoli considered the thermionic emission as a ionisation process and applied Saha's theory of equilibrium ionisation in order to obtain the electronic gas pressure as:

$$p_e = \frac{(2\pi m_e)^{3/2}}{h^3} k^{5/2} T^{5/2} e^{-eV/kT} = \frac{(2\pi m_e k)^{3/2}}{h^3} T^{3/2} e^{-eV/kT} kT = nkT \quad (10.9)$$

For an electron gas in equilibrium with the metal array, the number of emitted electrons per unit time per unit surface was $N = n(kt/2\pi m)^{1/2}$, whence Richardson formula:

$$I = \frac{2\pi m_e e k^2}{h^3} T^2 e^{-eV/kT} \quad (10.10)$$

with Dushman factor $b = eV/k$.

Pontremoli criticised Richardson's proposal to consider the photoelectric effect a case of thermionic emission [36]. Assuming Richardson's theory as true, with purely thermodynamic considerations, Pontremoli calculated the current density of the photoelectric emission corresponding to a quantity of emitted charge per energy

incompatible with the experimental results.⁹ Richardson's theory was wrong and the photoelectric and thermionic emissions had to be considered two different phenomena.

The last work done by Pontremoli in Rome was a theoretical quantum study of electric conductivity of flames containing alkaline salts [35]. He obtained the constant parameters in Wilson's formula $c = ((l^2 - 1) / l) (al + b)$, where c is the concentration by volume of the salt, l is the flame conductivity (with $l = 1$ when $c = 0$):

$$a = (N_F^+) {}_0^2 \left(\frac{h^2}{2\pi mkT} \right)^{3/2} e^{-eV/kT} \quad (10.11)$$

$$b = (N_F^+) {}_0 \quad (10.12)$$

where V is the ionisation potential, $(N_F^+) {}_0$ is the concentration by volume of positive ions in the flame. With these formulas, Pontremoli calculated the ratios b/a , which were compatible with the experimental values.

During the years he worked in Rome as Corbino's assistant, Pontremoli showed his abilities not only as a scientist but also as an organiser. He was one of the most active supporters of the Italian University Federation, whose promoting committee had his grandfather as one its members. The Federation sent Pontremoli to Prague as the Italian delegate to the international congress of the students (23 March–6 April 1921) called there by the Confédération internationale des étudiants. The congress assembly unanimously authorised the Confédération to admit Italy as a titular member thanks to Pontremoli's diplomatic activity. In 1924, once again his grandfather employed Pontremoli's ability in managing difficult enterprises: he was named secretary of the Italian Committee for the Transportation and the Settlement of the Armenian Refugees who had escaped the genocide in Anatolia, a committee who had his grandfather as co-president.

10.4 Pontremoli's Researches in Milan

On the 9th of December 1924 Pontremoli officially became lecturer ("libera docenza") of Superior Physics and was called to Milan to establish the Institute of Complementary Physics of the newly founded University of Milan. All he needed for the laboratories, the workshop, the library, the classrooms and the offices was bought only in part with the insufficient University funds. As the grandson of its founder and first president, Pontremoli convinced the Banca Popolare di Milano to help the University in financing the establishment of the radiology laboratory for researches on radioactivity, X-rays and UV-rays (see Sect. 10.5).

⁹E.g., for platinum, the theoretical value was 4.45×10^{-4} C/erg to be compared with the experimental value of 7.20×10^{-13} C/erg.

Pontremoli's assistants were: in 1924–1927 Glauco De Mottoni y Palacios (1901–1988), in 1925–1928 Enzo Pugno Vanoni (1899–1939), in 1927–1928 Maria De Marco (1896–1941), and informally in 1927–1928 Giorgio Diaz De Santillana (1902–1974).

De Mottoni graduated in Electrotechnical Engineering in Milan in 1924 and in Applied Mathematics in 1926. He was Pontremoli's assistant in 1924–1927 and carried on researches in optics. He had also been working as an astronomer, which would have been his scientific career thereafter. Pugno Vanoni [52] graduated in Electrotechnical Engineering in Milan in 1922. In 1925 he was called by the University of Milan to teach Electrotechnics and Physics of Röntgen Radiations, he became Pontremoli's assistant and taught Experimental Physics to the students of medicine. After Pontremoli's death he supplied him in the course of Experimental Physics at the Faculty of Sciences and was *pro tempore* director of the Institute of Complementary Physics. Pugno Vanoni planned all the electrotechnical systems of the Institute of Complementary Physics (see [52] and [53]). In 1931 he became professor of Electrotechnics at the University of Padua where he prematurely died. De Marco graduated in Civil Engineering in Rome in 1927 and was, most probably, the first female electrotechnician in Italy. She moved back to Rome for family reasons in 1930. De Santillana [57] graduated in Physics in Rome in 1925. He was Pontremoli's assistant for a short time (1927–1928) before going back to Rome and becoming a historian of science. In 1936 he moved to the United States and became in 1942 professor of history and philosophy of science at the M.I.T.

In 1926 Pontremoli was one of three winners of the public competition for a chair of Theoretical Physics (see Sect. 10.6). Due to some bureaucratic problems, he was appointed professor of Theoretical Physics only in April 1927. At the same time, the University started to plan the transfer of the Institute to the Science Palace. This would have meant to rebuild everything after so few years. Both these events, together with the death of his grandfather on the 29th of March deeply sadden Pontremoli. To recover from the stroke, Pontremoli worked extensively in the organisation of the International Congress of Physics held in Como. Also the choice to join the polar expedition in 1928 was a way to start again a more exciting life (see Sect. 10.7).

Pontremoli's first researches in Milan concerned advanced classical optics. In the first place [8, 37], Pontremoli studied the orientation of dielectric and diamagnetic ellipsoidal molecules of a gas in an external electric or magnetic field. He found the distribution of the number dN of molecules whose polar axis is at an angle between θ and $\theta + d\theta$ from the external field H :

$$dN = N \frac{e^{H^2(P-E)\cos^2\theta/2kT} \sin\theta d\theta}{\int_0^\pi e^{H^2(P-E)\cos^2\theta/2kT} \sin\theta d\theta} \quad (10.13)$$

where $P \cos\theta$ and $E \sin\theta$ were the molecules momentum along the polar resp. normal axis, with $P > E$. If $H = 0$, the distribution was isotropic with $dN_0 = N/4\pi d\omega = N/2 \sin\theta d\theta$. Pontremoli wondered if there was a particular direction with $dN = dN_0$. By equating the two expressions, he obtained the equation

$$e^{-a \sin^2 \theta} (1 + 2a \cos^2 \theta) = 1 \quad (10.14)$$

For usual experimental conditions $a < 10^{-4}$, so that the solution is given by $\cos \theta = \pm \sqrt{1/3}$. The statistical distribution of the polar axis of the molecules was therefore invariant on the surface of a $54^\circ 44'7''$ -wide cone whose axis was the direction of the orientating field. This result was later applied to a birefracting material [38].

With De Mottoni, he then studied the diffusion of light inside a material in an external, constant electric or magnetic field [7]. He considered again the case of dielectric, diamagnetic, polarisable, ellipsoidal molecules. Starting from the equations of motion for an electron in an external field, he found the polar and equatorial polarisation coefficients for a molecule. The consequent potential energy was integrated in the Boltzmann distribution for the molecules within a given conic solid angle. With this distribution, Pontremoli obtained the intensity of the diffused radiation in the three cartesian directions.

In 1926, Pontremoli studied again the birefracton in moving colloidal materials, the subject of his graduation thesis. Dario Graffi (1905–1990) had made some experiments on the birefracton of Bravais colloidal iron in motion inside a tube with rectangular section. Pontremoli tried to apply his theory to Graffi's results [39] but he was not even able to replicate them experimentally. His theory was valid for a liquid with constant speeds of dilation and flux in all the points on a same axis parallel to the observation axis. At the same time, Graffi was able to observe only the first family of fringes. In Graffi's opinion Pontremoli's theory unified the effects of two distinct kinds of birefracton, one due to the orientation of the molecules, and the other due to the internal deformation of the liquid. Further experiments by Graffi showed that Pontremoli's theory correctly foresaw the orientation of the particles and that other discrepancies were due to the dichroism of the liquid.

Another subject he had already studied in Rome was spectroscopy. In Milan, he developed a procedure to calculate how long is the monochromatic emission for a transition between two stationary circular orbits (in the framework of Bohr's atomic theory) [40]. He considered a hydrogen-like atom (with $Ze = E$), whose electron moved from the fundamental orbit (with radius r_1) to the second stationary orbit (r_2) under the action of the Coulombian attraction and of a fictional force describing the energy loss due to classical irradiation. Pontremoli's formula for the emission duration was:

$$t_{r_1 \rightarrow r_2} = \frac{m_e^2 c^3}{4e^3 E} (r_2^3 - r_1^3) \quad (10.15)$$

The experimental values for the first line of the Balmer, Lyman, and Paschen series were compatible with the values calculated by Pontremoli. He then wrote the orbital radiuses as a function of the quantum number n and re-wrote his formula as:

$$t_{r_1 \rightarrow r_2} = \frac{m_e c^3}{64\pi^2 e^2} \frac{1}{r^2 z^4} (n_2^6 - n_1^6) \quad (10.16)$$

Now Pontremoli could write his formula in a third way, using the quantum number n to express the orbital periods:

$$t_{r_1 \rightarrow r_2} = \frac{3m_e c^3}{16\pi^2 e^2} (T_2^2 - T_1^2) \quad (10.17)$$

Pontremoli developed then an approximate formula for transitions between two stationary orbits with large quantum numbers:

$$t = \frac{3m_e c^3}{8\pi^2 e^2} \left(\frac{1}{2RZ^2} \right)^{1/3} \frac{1}{\nu^{5/3}} \quad (10.18)$$

where ν is the frequency of the radiation emitted by the electron. Eventually, Pontremoli proved that, for large quantum numbers n and circular orbits, the mean life of an excited state corresponded to the duration of emission from that state.

In a further work [42], Pontremoli applied his formula to the K_α emission line in the X-ray spectrum, with the orbit radiuses given by Moseley's law. He obtained a formula (with the numerical values of all physical constants) as a function of Z :

$$t = \frac{9.94 \times 10^{-10}}{(Z - 1)^4} \text{ s} \quad (10.19)$$

or as a function of ν :

$$t = \frac{5.99 \times 10^{21}}{\nu^2} \text{ s} \quad (10.20)$$

The immediate consequence was the following law: the product of the emission time and the squared emission frequency of a K_α line is a constant for any chemical element. Pontremoli's law permitted to calculate in a very easy way the emission time once known the K_α line frequency from sodium (9.94×10^{-14} s) to uranium (1.45×10^{-17} s).

As a last subject of research, we can mention that in 1927 Pontremoli attempted a theoretical study of electric and thermal conductivity of metals [47]. A quantum study of metals was still far from having a satisfying formulation. While it was possible to calculate some coefficients when the metals were part of a saline structure, the same was not true when the structure was a non-ionic one. Pontremoli was well aware of the fact that the formulae he had derived were only approximations and that new experimental data were required to try a better formulation.

Pontremoli was also engaged in scientific divulgation. He wrote papers on astrophysics [41], the experimental foundations of quantum theory [44], the scientific progress [45], and the disintegration of matter [46]. For his colleagues of the Science Faculty he gave two seminary talks on some experiments with electric currents [48] and quantum mechanics [49].

10.5 The Institute of Complementary Physics

The University planned to fund Pontremoli with 500,000 lire for the establishment of the Institute of Complementary Physics¹⁰ (150,000 lire in 1924–1925, 200,000 lire in 1925–1926, and 150,000 in the following years). These funds were meant in order to buy the scientific furniture, the books and journals for the library, and to build the electric plants and the workshop. Pontremoli was but aware that a modern Institute of Physics in the industrial capital of Italy needed more funds than those granted by the University. He got from the Banca Popolare di Milano 100,000 lire for the radiation laboratory, and from other private patrons and industries other funds or laboratory instruments.

The Institute location was in a school building in Via Sacchini 34. Pontremoli personally followed the works to adapt the building to a scientific institute. In a few months, the new Institute was ready to work both in teaching and research. The Institute was responsible for the Physics courses for the Faculty of Sciences and for the specialisation courses. Pontremoli himself taught Complementary Physics, Electrotechnics and Radiation Physics (for the students of Medicine), and Physiologic Optics (for the students of Ophthalmology). From 1926, he also taught Theoretical Physics.

The Institute occupied the first floor, the ground floor, and the underground floor of the building.

On the first floor, in the northern aisle, there were the research laboratories, the director's office, the library, the director's experimental room, a first assistants' experimental room, a first laboratory for graduation researches, the dark room for photography, the spectroscopy laboratory with a Corbino-Trabacchi device, a second assistants' experimental room, and a little scales room. In the southern aisle, there were a second laboratory for graduation researches, a dark laboratory for optics, two experimental rooms, the classroom for eighty students with a preparation room for classroom experiments. All movable instruments, when not used in the laboratories, were held in the corridor wardrobes.

As for the library, Pontremoli and his assistants bought the whole collection of *Zeitschrift für Physik* (which started in 1920) and of *Physikalische Zeitschrift* (in 1900). The other journals had been bought instead only from 1924 on. They were: *Annalen der Physik*, *Annales de Physique*, *Astrophysical Journal*, *Comptes Rendus de l'Academie des Sciences*, *Comptes Rendus de la Société Polonaise de Physique*, *General Electric Review*, *Journal de Physique*, *Journal of the American Institute of Electrical Engineering*, *Mount Wilson Contributions*, *Nature*, *Philosophical Magazine*, *Physical Review*, *Physikalische Berichte*, *Zeitschrift für technische Physik*. With the aim to have at their disposal the most important journals in their fields of research, they preferred to spend only a small part of the library funds in buying books.

¹⁰Information on the establishment of the Institute of Complementary Physics can be found in: [52] and [53].

On the ground floor, there were the workshop and the radiation laboratory with two experimental rooms and one corridor. The radiation laboratory was active in scientific and technical measurements on X-rays and could certificate the results of the tests and researches carried on: tests of radiology devices, X-ray tubes, ionic and thermoelectric valves; calibration of devices; X-ray spectrographic measurements; ionisation measurements. The radiation laboratory was planned to be the fundamental core of the Institute and of a planned institute for the cure of cancer in Milan. The Technical Commission of the Italian Society of Medical Radiology entrusted the radiation laboratory in Milan and the physics laboratory of the Public Health Institute in Rome of all radiological studies, tests and checks on account of the same Society.

The radiation laboratory was divided into four working sections: electrotechnics, testing of tubes and thermoelectric valves, calibration of radiometric devices, and research. The research section was mostly involved in radio-metalloscopy experiments to show defects in the metal structures, the study of the energy distribution emitted by tubes supplied by different tension shapes, and the study of radiation-absorbing screens which minimised the diffused radiation. They also planned to study the rules concerning the tests on radiological devices.

On the underground floor there were the electric plants, the batteries storage room, and two laboratory rooms where they could carry on researches which need not to be affected by mechanical vibrations from the concrete structures of the building. The electric plants were composed of a Marelli 10.7 kW, 165/230 V dynamo group; a Giampiero Clerici & Co. 5 kVA three-phase 160/10,000 transformer; four 95 Ah Hensemberger batteries.

The instruments in the Institute were bought in a given order of importance according to the funds granted each year. In a first time they made built the fixed plants necessary for teaching and everyday research. The second step was the acquisition of the devices for the researches and of those which it was unpredictable that they could buy once the starting funds had been over.

The most important instruments were: an intensive Corbino-Trabacchi apparatus¹¹ with a 280,000 V rectified tension in the spectroscopy room; three big Littrow spectrograph¹² (one of quartz, and two of glass) for measurements with prisms¹³ in the visible and the ultraviolet; a spectrograph with constant deviation, with a Lummer-

¹¹A Corbino-Trabacchi apparatus was a device used to produce extremely-high tension electrical currents as power suppliers for X-ray tubes, see [6].

¹²The Littrow spectrograph is a prism with 30°, 60°, and 90° angles. The face opposing the 60° angle is coated with a reflective film.

¹³If a light beam enters a Littrow spectrograph at the Brewster angle, it undergoes the maximum dispersion. The linear dispersion observed on the screen was of 1.14 m in the 2000–8000 Å range with the quartz prism, and of 0.70 m in the 4000–8000 Å range with the two glass prisms.

Gehrcke plate,¹⁴ a Michelson grating,¹⁵ a Fabry-Pérot etalon,¹⁶ and a set of oculars; a monochromator for the infrared, visible and ultraviolet radiations, calibrated in the 1850–100,000 Å range; a big Möll microphotometer¹⁷ on free temporary loan by Kipp and Zonen¹⁸; a Yvon spectrophotometer for the studies of absorption, diffusion, polarimetry, photometry, homochromatic and heterochromatic spectrophotometry, and position microphotometry; a big circle of Jamin and Sènarmon¹⁹ to study refractions and reflections; a Jamin optical bench to study diffractions and interferences; a Weiss electromagnet²⁰ to produce magnetic fields up to 38,000 gauss; a high-sensitivity Compton electrometer²¹; a Hartmann and Braun millivoltmeter and milliamperometer.

After Pontremoli's death, Giovanni Polvani was called to Milan as professor of Experimental Physics. The Institute was moved to the new Science Palace, located in Via Saldini, in the building hosting the institutes of Mathematics and of General Industrial Chemistry. The library and the instruments were moved to the new Institute, the laboratories and rooms were dismantled: Pontremoli's Institute therefore came abruptly to an end.

10.6 The Public Competition for the Chair of Theoretical Physics

The first public competition for three chairs of Theoretical Physics in Rome, Florence, and Milan, was published in May 1926. With Theoretical Physics they explicitly meant quantum theoretical physics, a subject which was not taught at that time in Italian universities. Some young Italian physicists had studied the early quantum theories, atomic or nuclear quantum phenomena, and the very recent quantum mechanics. Fermi had already invented a new quantum statistics to study identical

¹⁴The Lummer-Gehrcke plate was a high-resolution, multiple-beam interferometric device. It was made of a plane-parallel quartz or glass plate. A light ray entering it undergoes multiple reflections inside the plate; the refracted, parallel, emerging rays can made interfere in the focal plane of a converging lens. The Lummer-Gehrcke plate was later replaced by the Fabry-Pérot etalons.

¹⁵A typical Michelson grating has 500 rulings per millimeter.

¹⁶A Fabry-Pérot etalon is a couple of two parallel, facing, partially-reflecting mirrors. They are the main component of the Fabry-Pérot interferometer.

¹⁷The Möll microphotometer used: a thermocouple with galvanometer as a receiver; a slit to reduce the area on the photographic plate illuminated by the deflected beam to the exact dimensions of the incident beam. The incident and the deflected beams were connected by a linear relation.

¹⁸For the microphotographs of X-rays produced by Pontremoli with a Coolidge tube, see [43].

¹⁹The circle of Jamin and Sènarmon was an orientable brass circle. Two movable small circles, perpendicular to the large one, carried on them the analyser and the polariser connected to alidades and verniers for angle measurements.

²⁰The Weiss electromagnet is the historical model for most modern electromagnets, with cylindrical iron-cobalt poles wrapped by copper tubes as coils.

²¹The Compton electrometer was a quadrant electrometer used for radioactivity measurements.

particles with a half-integer spin. The three chairs of Theoretical Physics would have therefore been a further step in shifting Italy out of a peripheral context. Corbino made use of his political power to have the public competition and to start it as soon as possible. His main aim was to have Fermi as professor of Theoretical Physics in Rome.

The commission instructed to evaluate the curriculum vitae and the publications of the candidates was formed by Antonio Garbasso (1871–1933) as president, Corbino as relator, Gian Antonio Maggi (1856–1937), Michele Cantone (1857–1932), and Quirino Majorana (1871–1957). Corbino, Garbasso, and Majorana were educated in classical physics but they also made some researches in modern physics: Corbino studied the Zeeman effect and carried on many researches on the X-rays; Garbasso studied spectroscopy and the nature of X-rays, and gave a theoretical explanation of the Stark-Lo Surdo effect; Majorana made experiments to test that the light speed is constant (he was an anti-relativist), carried on researches in photoelectricity and on the nature of X-rays. Cantone and Maggi had instead studied only classical physics; they most probably were commissioners only in order to ensure the legal number of five members.

Four candidates took part to the competition: Carlo Del Lungo (1867–1950), Fermi, Persico, and Pontremoli. Del Lungo was an expert in classical thermodynamics, he had not carried on any research, nor experimental nor theoretical, in modern physics. He was therefore considered not suitable to win the competition.

Fermi was easily considered the best one:

The commission is therefore unanimous in declaring that prof. Fermi is highly deserving of covering the chair of Theoretical Physics under competition, and believes it can rest on him the best hopes for the affirmation and the future development of theoretical physics in Italy.²²

As for Pontremoli, the commission took into consideration not only his education and scientific activity in contemporary physics, but also his ability as an organiser:

The commission, after an extensive examination of the candidate's work, expresses collegially and unanimously the following opinion: Prof. Pontremoli demonstrates with his titles a wide and diverse culture, also in the fields of modern theoretical physics; an excellent mathematical education, a lively wit, a fervid fantasy, and a great passion for scientific research.

[The commission] has also well-established his laudable didactic attitudes and particularly the merits he acquired in creating in a short time in Milan a Physics Institute that is among the best equipped in Italy for some modern subjects of physics research. Of course, if he had not been found himself involved in this task he has so praisingly absolved, his scientific production in the last years would have been, if not more abundant, more valuable and profound, and therefore less subject to some reserves. Therefore, the commission is unanimous in recognising that Prof. Pontremoli, for his scientific and didactic maturity, is well worth covering a chair of Theoretical Physics, as well as being very worth to cover the chair he is now in charge for.²³

²²Archivio di Stato di Roma, MPI. DGIS DIV 1, Conc. catt. univ. 1924–1954. Minutes of the Theoretical Physics competition in the Royal University of Rome, 2nd meeting, pp. 3–4.

²³Ibid., 3rd meeting, pp. 5–6.

The commission then voted for the three chairs. For the first chair Fermi got five votes out of five, for the second chair Persico got three votes and Pontremoli two votes, for the third chair Pontremoli got five votes. It was not easy to classify Persico and Pontremoli, who were more or less at the same level but with rather different abilities in facing theoretical physics:

It concurrently results from the discussion that the security and the degree of efficiency in the scientific inquiry, the profound culture in higher mathematics and high mechanics, the clear and inventive vividness of the mind prevail in Persico. Instead, in Pontremoli, the sparkling liveliness of the mind, the spirit of initiative, the adherence to the broad and sometimes turbid movement of modern physics prevail.²⁴

After the vote (the commissioners' names who voted in favour of Pontremoli are not reported in the minutes), the three winners were official declared in the following order: 1st Fermi unanimously; 2nd Persico by majority; 3rd Pontremoli unanimously. The establishment of the Institute of Physics with its entanglement between scientific practices and institutional abilities had played a fundamental role in Pontremoli's victory.

10.7 The Polar Expedition

In October 1927, Umberto Nobile (1885–1978) asked Pontremoli to join him in his second expedition to the Arctic on board of the *Italia* airship.²⁵ Italian polar expeditions were in the first place a matter of political prestige, which showed Italy as a leading country in airship and aircraft technology the world over. The 1928 *Italia* expedition, instead, should have been the first purely scientific polar expedition with the aim of study the physical and biological aspects of the arctic sea north to Scandinavia. Pontremoli was well aware of the immense potential offered by the aviation technology in the scientific research, given his experience as a drachen pilot during the First World War. He had already interested himself in the first polar expedition on board of the *Norge* airship. He thus accepted Nobile's proposal and started to organise the scientific aspects of the expedition. His scientific abilities and his passion for adventures and explorations would again put Italy further more away from a peripheral role in physics.

Two other scientists took part in the expedition: Finn Malmgren (1895–1928) and František Běhounek (1898–1973). Malmgren was a Swedish meteorologist and assistant professor at Otto Pettersson's hydrographic institute for oceanic studies in Bornö. He was a member of the 1922–1925 expedition on board of the *Maud* ship, led by Roald Amundsen and Harald Sverdrup, and of the 1926 expedition on board of the *Norge* airship. He was severely hurt during the *Italia* accident; his body was last seen on the ice pack almost two months later. Běhounek was a Czech physicist, trained in radiology by Marie Curie, an expert in radioactivity and cosmic radiation from

²⁴Ibid., pp. 6–7.

²⁵On the Italian polar expeditions, see [21].

the Czechoslovak Radiological Institute. Luigi Palazzo (1861–1933), director of the Central Institute for Meteorology and Geodynamics in Rome, only collaborated in the organisation of the expedition but did not take part in it.

Pontremoli planned the scientific laboratory [9] and the sets of measurements to be done in the following subjects: atmospheric electricity, gravimetry, geomagnetism, meteorology, oceanography, optics, penetrating radiation (cosmic rays), and the transmission of radio-waves in the Arctic atmosphere. As we can see, the scientific interest in polar researches affected both traditional subjects (geomagnetism, gravimetry, oceanography) in a still unknown environment, and new subjects (atmospheric radioactivity, penetrating radiation, radio-wave transmission).

Pontremoli planned the instruments which had to be small and light enough to be carried in the control cabin of an airship. In particular, he built the instruments for the first measurements ever of the geomagnetic (with Palazzo) and gravitational fields in the Arctic. Furthermore, Pontremoli tested all the devices in refrigerators in order to be sure of their proper working at low temperatures. Malmgren got the meteorological instruments, while Běhounek provided a Kolhörster electrometer for the penetrating radiation, an electrometer for atmospheric conductivity, metal plates for atmospheric radioactivity, and a Wiechert electrometer for the electric gradient in atmosphere. For the study of short wave transmissions, they used devices built by Philips and Allocchio and Bacchini. The latter firm also built the sounder for oceanographic research with inside a resistance thermometer for electrolytic conductivity and seawater temperature, and a Chauchard apparatus for seawater salinity.

The Italia airship started from Milan in the night between the 14th and the 15th of April 1928. On the 6th of May, it landed in the King's Bay (Svalbard Islands), which was the base camp for the whole expedition. During the third flight, on the 25th of May, due to the sudden and fast loss of altitude, the Italia airship hit the ice pack and bounced back into the air. During the collision, ten men fell on the ice pack. Only eight of them would have been later rescued alive. The drifting airship most probably sank into the Sea of Barents. Pontremoli was on board with other five people. Some of the preparatory studies and the very few surviving data were later published in a volume edited by Nobile (see [1, 9, 22, 24–26]), with an incomplete testimony of Pontremoli's scientific researches in that adventurous expedition since his scientific notebook had disappeared with him.

References

1. F. Běhounek, Forschungen über atmosphärische Elektrizität, in *Die Vorbereitungen und die wissenschaftlichen Ergebnisse der Polarexpedition der "Italia"*, ed. by U. Nobile, Petermanns Mitteilungen, Ergänzungsheft Nr. 205 (1929), pp. 46–62
2. L. Belloni, Giovanni Polvani e l'Istituto di Milano. *Il Nuovo Saggiatore* **4**, 35–49 (1988)
3. T. Calzecchi Onesti, Sulla conduttività delle limature metalliche. *Nuovo Cimento* **3**, **XIII**, 58–64 (1884)

4. T. Calzecchi Onesti, Sulla conduttività delle limature metalliche. *Nuovo Cimento* 3, **XVII**, 38–47 (1885)
5. T. Calzecchi Onesti, Di una forma che può darsi all'avvisatore microsismico. *Nuovo Cimento* 3, **XIX**, 24–26 (1886)
6. O.M. Corbino, G. Trabacchi, Dispositivi per produrre correnti di altissima tensione e di senso costante atte all'alimentazione dei tubi per raggi X. *Il Nuovo Cimento* **XI**(1), 361–376 (1916)
7. G. De Mottoni, A. Pontremoli, La diffusione della luce in un mezzo sottoposto ad un campo elettrico o magnetico costante. *Rendiconti Istituto Lombardo Scienze Lettere* **LVIII**, 801–807 (1925)
8. G. De Mottoni, A. Pontremoli, Un caso particolare di diffusione della luce in presenza di un campo orientante costante. *Rendiconti Istituto Lombardo Scienze Lettere* **LIX**, 60–61 (1926)
9. G. De Mottoni, E. Pugno Vanoni, Einige Instrumente die von prof. Aldo Pontremoli für die Polfahrt Nobiles im Jahre 1828 konstruirt worden sind, in *Die Vorbereitungen und die wissenschaftlichen Ergebnisse der Polarexpedition der "Italia"*, ed. by U. Nobile, Petermanns Mitteilungen, *Ergänzungsheft* Nr. 205 (1929), pp. 89–98
10. S. Dushman, Electron Emission from metals as a function of temperature. *Phys. Rev.* **21**, 623–636 (1923)
11. E. Fermi, A. Pontremoli, Sulla massa della radiazione in uno spazio vuoto. *Atti Accademia Lincei Rendiconti* **XXXII**, 162–164 (1923)
12. B. Finzi, Aldo Pontremoli. *Atti Seminario Matematico e Fisico di Milano* **2**, xi–xii (1928)
13. L. Flamm, Die charakteristischen Maßzahlen für das Elektron in ihrer Verknüpfung mit den Strahlungskonstanten. *Physikalische Zeitschrift* **18**, 515–521 (1917)
14. L. Gariboldi, Pontremoli, Aldo. *Dizionario Biografico degli Italiani*, vol. 84. Treccani, Roma (2015). [http://www.treccani.it/enciclopedia/aldo-pontremoli_\(Dizionario-Biografico\)/](http://www.treccani.it/enciclopedia/aldo-pontremoli_(Dizionario-Biografico)/)
15. L. Gariboldi, Aldo Pontremoli e l'Istituto di Fisica di Milano, in *Intorno a Galileo: la storia della fisica e il punto di svolta galileiano*, ed. by E. Giannetto, E. Giannini, M. Toscano (Bergamo, 2011), pp. 211–217
16. L. Gariboldi, La nascita e i primi sviluppi degli studi di fisica. *Annali di storia della università italiane* **11**, 261–276 (2007)
17. L. Gariboldi, The neutron before the neutron: Pontremoli's compound models, in *Atti del XXXV Convegno annuale/Società Italiana degli Storici della Fisica e dell'Astronomia*, ed. by S. Esposito (2015), pp. 267–271
18. G.P. Giordana, *Vita di Aldo Pontremoli* (Formiggini, Roma, 1933)
19. J.L. Glasson, Attempts to detect the presence of neutrons in a discharge tube. *Philos. Mag.* **42**, 892–904 (1921)
20. G.J. Holton, *The Scientific Imagination: With a New Introduction* (Harvard University Press, Cambridge MA, 1998)
21. L.H. Juvik, Italienske reiser på Nordkalotten før 1945, Tromsø (2003)
22. F. Malmgren, Bericht über den Flug nach Nordland (Nikolaus-II Land), in *Die Vorbereitungen und die wissenschaftlichen Ergebnisse der Polarexpedition der "Italia"*, ed. by U. Nobile, Petermanns Mitteilungen, *Ergänzungsheft* Nr. 205 (1929), pp. 63–65
23. A.L. McAulay, The recoil of hydrogen nuclei from swift alpha particles. *Philos. Mag.* **6**(42), 892–904 (1921)
24. U. Nobile, Aerometeorologische Beobachtungen an der Königs-Bucht im April und Mai 1928, in *Die Vorbereitungen und die wissenschaftlichen Ergebnisse der Polarexpedition der "Italia"*, ed. by U. Nobile, Petermanns Mitteilungen, *Ergänzungsheft* Nr. 205 (1929), pp. 66–79
25. U. Nobile, Das geographische Problem des Arktis, das Programm und die Durchführung der Flüge der "Italia", in *Die Vorbereitungen und die wissenschaftlichen Ergebnisse der Polarexpedition der "Italia"*, ed. by U. Nobile, Petermanns Mitteilungen, *Ergänzungsheft* Nr. 205 (1929), pp. 9–45
26. L. Palazzo, Vorstudien für die Erdmagnetischen Forschungen auf der Luftschiffexpedition in die Arktis, in *Die Vorbereitungen und die wissenschaftlichen Ergebnisse der Polarexpedition der "Italia"*, ed. by U. Nobile, Petermanns Mitteilungen, *Ergänzungsheft* Nr. 205 (1929), pp. 80–88

27. A. Pontremoli, Sul luogo fisico delle frangie nella doppia rifrazione accidentale meccanica di un liquido in moto piano permanente. *Atti Accademia Lincei Rendiconti* **XXX**, 216–220 (1921)
28. A. Pontremoli, La doppia rifrazione accidentale meccanica nei liquidi. *Atti Accademia Lincei Memorie* **XIII**, 593–616 (1922)
29. A. Pontremoli, Le equazioni di propagazione di Maxwell per un dielettrico sottoposto ad un campo elettrico e magnetico longitudinale. *Atti Accademia Lincei Rendiconti* **XXXI**, 189–192 (1922)
30. A. Pontremoli, Potere rotatorio creato in un mezzo isotropo a molecole simmetriche da un campo elettrico e magnetico longitudinali e costanti. *Atti Accademia Lincei Rendiconti* **XXXI**, 434–440 (1922)
31. A. Pontremoli, Sul neutrone di Rutherford. *Atti Accademia Lincei Rendiconti* **XXXII**, 277–280 (1923)
32. A. Pontremoli, Un nuovo effetto del campo magnetico sulla scarica nei gas rarefatti. *Atti Accademia Lincei Rendiconti* **XXXII**, 158–161 (1923)
33. A. Pontremoli, Sulla emissione termoionica. *Atti Accademia Lincei Rendiconti* **XXXII**, 211–214 (1923)
34. A. Pontremoli, Sulla scarica nei gas rarefatti. *Nuovo Cimento* **XXVI**, 115–121 (1923)
35. A. Pontremoli, Sulla conducibilità elettrica delle fiamme contenenti sali alcalini. *Atti Accademia Lincei Rendiconti* **XXXIII**, 223–227 (1924)
36. A. Pontremoli, Sull'effetto fotoelettrico. *Rendiconti Istituto Lombardo Scienze Lettere* **LVII**, 567–574 (1924)
37. A. Pontremoli, Sull'orientamento in un campo elettrico o magnetico costante nella ipotesi dell'anisotropia molecolare. *Atti Accademia Lincei Rendiconti* **II**, 328–331 (1925)
38. A. Pontremoli, Una esperienza caratteristica sulla birifrangenza elettrica o magnetica. *Atti Accademia Lincei Rendiconti* **II**, 416–419 (1925)
39. A. Pontremoli, Circa alcune nuove ricerche sulla birifrangenza accidentale dei colloidi in movimento. *Atti Accademia Lincei Rendiconti* **III**, 75–77 (1926)
40. A. Pontremoli, Sulla durata di emissione delle radiazioni monocromatiche e la vita media degli stati stazionari. *Atti Accademia Lincei Rendiconti* **III**, 149–154 (1926)
41. A. Pontremoli, Atomi ed astri. *Periodico di Matematiche* **4**(6), 99–111 (1926)
42. A. Pontremoli, Sulla durata di emissione della riga K_{α} pei vari elementi. *Rendiconti Istituto Lombardo Scienze Lettere* **LIX**, 340–342 (1926)
43. A. Pontremoli, Sul funzionamento del tubo Coolidge in corrente continua o sinusoidale rad-drizzata e le conseguenti caratteristiche dell'emissione dello spettro continuo dei raggi X. *Elettrotecnica* **XIII**, 649–650 (1926)
44. A. Pontremoli, Le basi sperimentali della teoria dei quanti. *Elettrotecnica* **XIII**, 793–796 (1926)
45. A. Pontremoli, Le scuole scientifiche ed il progresso della scienza. *Fiera Letteraria* (1926)
46. A. Pontremoli, La disintegrazione della materia. *Fiera Letteraria* (1926)
47. A. Pontremoli, Sulla conducibilità elettrica e termica dei metalli. *Rendiconti Istituto Lombardo Scienze Lettere* **LX**, 273–280 (1927)
48. A. Pontremoli, Esperienze su correnti piane parallele incontranti un cilindro circolare. *Atti Seminario Matematico e Fisico di Milano* (1927)
49. A. Pontremoli, Le nuove vie della meccanica. *Atti Seminario Matematico e Fisico di Milano* (1927)
50. E. Pugno Vanoni, Laboratorio radiologico di taratura e ricerca **XIII**, 697–702 (1926)
51. E. Pugno Vanoni, Aldo Pontremoli. *Nuovo Cimento* **VII**, 41–49 (1930)
52. E. Pugno Vanoni, Il Nuovo Cimento **XVI**, 333–336 (1939)
53. Regia Università di Milano: Istituto di Fisica Complementare, in *Annuario 1926–1927*
54. Regia Università di Milano: In memoria del Prof. Aldo Pontremoli, in *Annuario 1929–1930*, pp. 153–163
55. E. Rutherford, Nuclear constitution of atoms. *Proc. R. Soc. A* **97**, 374–400 (1920)
56. L. Silberstein, Über die gegenseitige Masse kugelförmiger Elektronen. *Physikalische Zeitschrift* **12**, 87–91 (1911)

57. N. Sivin, Éloge: Giorgio Diaz de Santillana, 1902–1974. *Isis* **67**, 439–443 (1976)
58. A. Sommerfeld, *Atombau und Spektrallinien* (Vieweg, Braunschweig, 1922)
59. G. Tagliaferri, Le scienze esatte all'Università di Milano, in *Storia di Milano*, vol. 18 (Istituto dell'Enciclopedia Italiana, Roma, 1995), pp. 659–677
60. R.W. Wood, An extension of the balmer series of hydrogen and spectroscopic phenomena of very long vacuum tubes. *Proc. R. Soc.* **A97**, 455–470 (1920)
61. R.W. Wood, Hydrogen spectra from long vacuum tubes. *Philos. Mag.* **42**, 729–745 (1921)
62. R.W. Wood, Atomic hydrogen and the balmer series spectrum. *Philos. Mag.* **44**, 538–546 (1922)
63. R.W. Wood, Spontaneous incandescence of substances in atomic hydrogen gas. *Proc. R. Soc.* **A102**, 1–9 (1922)

Chapter 11

The Role of Monitoring Time and Detectors Efficiencies in Time-Continuous Quantum Magnetometry



Francesco Albarelli, Matteo A. C. Rossi, Matteo G. A. Paris
and Marco G. Genoni

Abstract We consider the estimation of a weak magnetic field B acting on a continuously monitored ensemble of atoms subjected to collective transverse noise. If N atoms are prepared in a coherent spin state and are not continuously monitored, the estimation precision scales with the total number of atoms according to the standard quantum limit $\delta B^2 \sim 1/N$. Remarkably, time-continuous monitoring of light that is coupled with the atomic ensemble, allows to achieve a Heisenberg limited precision $\delta B^2 \sim 1/N^2$. However this is typically obtained only for a large enough number of atoms N and with an asymptotic constant factor depending on the parameters characterizing the experiment. In this proceeding, after reviewing the analytical derivation of the effective quantum Fisher information that quantifies the ultimate precision achievable, we specifically address the role played by monitoring time and detectors measurement efficiency in obtaining a Heisenberg limited scaling. In particular we analyze the dependence on these experimentally relevant parameters of the asymptotic constant factor characterizing the effective quantum Fisher information, and, more importantly, the minimum value of atoms needed to observe the desired quantum enhancement.

11.1 Introduction

The aim of quantum metrology is to take advantage of purely quantum-mechanical features to accurately estimate physical parameters, with a better precision than classical schemes [1, 2]. This framework has been fruitfully applied to magnetometry, in particular to the estimation of the intensity of a magnetic field B (with known

F. Albarelli · M. A. C. Rossi · M. G. A. Paris · M. G. Genoni (✉)
Quantum Technology Lab Dipartimento di Fisica “Aldo Pontremoli”,
Università Degli Studi di Milano, 20133 Milan, Italy
e-mail: marco.genoni@fisica.unimi.it

direction) probed by an ensemble of N two-level atoms with total spin $J = N/2$. This is the metrological task we will consider in the present proceeding.

The standard quantum metrological scheme goes as follows: an initial state of the probing system is prepared, the system then undergoes an evolution governed by the magnetic field and finally the probe is measured. In the noiseless unitary case the evolution simply amounts to a rotation of the state vector on the Bloch sphere [3–9]. The statistics produced by the final measurement is then processed to estimate B ; the most commonly used figure of merit to quantify the precision of the estimate is the mean-square error (MSE).

One of the most important features of these schemes is the scaling of the precision with the number N of atoms. The $1/N$ scaling of the MSE is the so-called standard quantum limit (SQL) to precision, which is obtained by preparing the atomic ensemble in a separable spin coherent state. On the other hand a $1/N^2$ scaling denotes the Heisenberg limit to precision and it can be reached by employing an entangled initial state, such as a GHZ or a spin squeezed state [10, 11]. Unfortunately, the Heisenberg scaling is easily lost in the presence of external noise [12–14] and a plethora of strategies to circumvent this limitation have been proposed [15–26].

We analyze a different metrological protocol, based on the continuous monitoring of the system in time, done by measuring the degrees of freedom of the environment [27, 28]. Remarkably, this approach does not rely on the preparation of an initial entangled state of the probe system. The main idea is that the time signal obtained from continuous monitoring can be employed to accurately estimate B . At the same time, the measurement back-action due to the monitoring squeezes the state and makes it more useful for magnetic field estimation. This protocol also includes a final strong measurement on the atomic system, analogously to the traditional schemes.

Similar ideas for quantum magnetometry with continuous measurements have been proposed and studied [29–34] and the topic has been explored in the context of quantum metrology more in general [35–42]. In [43] we rigorously analyzed the performance of this magnetometry scheme and we provided an appropriate Cramr-Rao bound which takes into account the information due both to the monitoring and to the final measurement. We concluded that Heisenberg scaling can be achieved, for a sufficiently long monitoring time and a sufficiently high number of atoms, while also showing that non unit efficiency of the detectors does not undermine this result. In the present proceeding we quantitatively characterize the parameter regimes which allow to obtain Heisenberg scaling, which is possible due to the analytic result obtained in [43]. In particular we study how the minimal number of atoms needed to observe Heisenberg scaling behaves as a function of the monitoring time and of the detector efficiency.

The manuscript is organized as follows. In Sect. 11.2 we review the physics of an atomic ensemble continuously monitored by measuring light after the interaction with the atoms. In Sect. 11.3 we introduce the metrological scheme and present

the appropriate Cramr-Rao bound, we then present the quantitative analysis of the minimal number of atoms needed to observe Heisenberg scaling. Section 11.4 concludes the paper with some final remarks.

11.2 Continuously Monitoring of an Atomic Ensemble

Let us consider an ensemble of N two-level atoms that can be described as a system with total spin $J = N/2$ with collective spin operators

$$\hat{J}_\alpha = \frac{1}{2} \sum_{i=0}^N \sigma_{i\alpha}, \quad \alpha = x, y, z, \quad (11.1)$$

where $\sigma_{i\alpha}$ denotes the Pauli matrices acting on the i -th spin (atom). The collective operators obey the commutation rules $[\hat{J}_i, \hat{J}_j] = i\varepsilon_{ijk} \hat{J}_k$ for angular momentum, where ε_{ijk} is the Levi-Civita symbol and we choose units such that $\hbar = 1$. We then assume that an unknown static and constant magnetic field B is acting on the ensemble, such that the unitary evolution is described by the Hamiltonian $\hat{H}_b = \gamma B \hat{J}_y$, where γ represents the strength of the coupling with the magnetic field. The atoms are then coupled to an electromagnetic mode $a_i n(t)$ via an interaction Hamiltonian $\hat{H}_{\text{int}} = \sqrt{\kappa} \hat{J}_z (\hat{a}_i n(t) + \hat{a}_i^\dagger(t))$ [32, 44]. If the light modes are not measured, the evolution of the atoms is described by a collective transverse noise Markovian master equation

$$\frac{d\rho}{dt} = -i\gamma B [\hat{J}_y, \rho] + \kappa \mathcal{D}[\hat{J}_z]\rho, \quad (11.2)$$

where the noise generator is perpendicular to the direction of the magnetic field and ρ is the density operator describing the state of the atomic system. The ensemble is initially prepared in a spin coherent state directed in the positive x direction,

$$|\psi(0)\rangle = \bigotimes_{k=0}^N |+\rangle_k = |J, J\rangle_x, \quad (11.3)$$

i.e. in a tensor product of eigenstates $|+\rangle$ of the operator σ_x for each two-level atom.

The unconditional dynamics of the spin component $\langle \hat{J}_x \rangle$ is given by

$$\frac{d\langle \hat{J}_x(t) \rangle}{dt} = \gamma B \langle \hat{J}_z(t) \rangle - \frac{\kappa}{2} \langle \hat{J}_x(t) \rangle. \quad (11.4)$$

Notice that the the dissipative part of the dynamics causes the shrinking of the spin vector $\langle \hat{\mathbf{J}} \rangle$, while the unitary part is responsible for its Larmor precession around the

y-axis. In this paper we assume to estimate *small* magnetic fields, such that $\gamma Bt \ll 1$, the solution of (11.4) can be approximated as

$$\langle \hat{J}_x(t) \rangle \approx \langle \hat{J}_x(0) \rangle e^{-\kappa t/2} = J e^{-\kappa t/2}. \quad (11.5)$$

Let us now consider a measurement scheme where we continuously monitor the light modes via homodyne measurements at the appropriate phase, immediately after the interaction with the ensemble, with a certain monitoring efficiency η . This in fact allows a continuous “weak” measurement of the collective spin operator \hat{J}_z and the evolution of the spin system is described by the following stochastic master equation

$$\begin{aligned} d\rho^{(c)} = & -i\gamma B[\hat{J}_y, \rho^{(c)}]dt + \kappa \mathcal{D}[\hat{J}_z]\rho^{(c)}dt \\ & + \sqrt{\eta\kappa} \mathcal{H}[\hat{J}_z]\rho^{(c)}dw_t, \end{aligned}$$

where dw_t denotes a Wiener increment [28]. We can express the measurement result at time t in terms of the infinitesimal photocurrent

$$dy_t = 2\sqrt{\eta\kappa} \text{Tr}[\rho^{(c)} \hat{J}_z]dt + dw_t. \quad (11.6)$$

Let us here remark that the collective noise present in the master equation (11.2) is suitable for describing the dynamics also in experimental setups where there is no engineering of the coupling to the atomic ensemble, with the purpose of performing continuous monitoring [45–47]. In this respect, assuming non-unit efficiency η may account for both: (i) homodyne detectors that can not capture all the photons that have interacted with the spin, (ii) environmental degrees of freedom that cannot be measured during the experiment and that induce the same type of noisy dynamics.

In the limit of a large number of atoms $N \gg 1$, the dynamics can be effectively described with the Gaussian formalism. This holds true as long as $\langle \hat{J}_x(t) \rangle \approx J$, i.e. $\kappa t \lesssim 1$.

Let us introduce the effective quadrature operators of the atomic sample, defined so that they satisfy the canonical commutation relation $[\hat{X}, \hat{P}] = i$, as [32, 33]

$$\hat{X} = \hat{J}_y/\sqrt{\bar{J}_t} \quad \hat{P} = \hat{J}_z/\sqrt{\bar{J}_t}, \quad (11.7)$$

where $\bar{J}_t \equiv |\langle \hat{J}_x(t) \rangle|$ (notice that in the limit of large spin J we can consider the unconditional average value $\langle \hat{J}_x(t) \rangle$, as the stochastic correction obtained via (11.6) would be negligible). The initial state $|\psi(0)\rangle$ corresponds to the vacuum state $(\hat{X} + i\hat{P})|0\rangle = |0\rangle$, which is Gaussian. Equation (11.6) now reads

$$\begin{aligned} d\rho^{(c)} = & -i\gamma B\sqrt{\bar{J}_t} [\hat{X}, \rho^{(c)}] dt + \kappa \bar{J}_t \mathcal{D}[\hat{P}]\rho^{(c)}dt + \\ & + \sqrt{\bar{J}_t\eta\kappa} \mathcal{H}[\hat{P}]\rho^{(c)}dw_t. \end{aligned}$$

Notice that (11.8) is quadratic in the canonical operators, and thus preserves the Gaussian character of states. The whole dynamics can thus be equivalently described in terms of the first and second moments only [48, 49]. More specifically, for estimating B , we only need the mean and the variance of the atomic momentum quadrature \hat{P} evaluated on the conditional state $\varrho^{(c)}$. We find, respectively,

$$d\langle \hat{P}(t) \rangle_c = -B\gamma \sqrt{J} e^{-\frac{\kappa t}{2}} dt + 2\text{Var}_c[\hat{P}(t)] \sqrt{\eta\kappa J} e^{-\frac{\kappa t}{2}} dw_t, \quad (11.8)$$

$$\frac{d\text{Var}_c[\hat{P}(t)]}{dt} = -4\eta\kappa J e^{-\frac{\kappa t}{2}} \left(\text{Var}_c[\hat{P}(t)] \right)^2. \quad (11.9)$$

We notice that (11.9) for the conditional variance is deterministic and can be solved exactly. If the initial state is a vacuum state, i.e. with $\text{Var}[\hat{P}(0)] = \frac{1}{2}$, we find the solution

$$\text{Var}_c[\hat{P}(t)] = \frac{1}{8\eta J \left(1 - e^{-\frac{\kappa t}{2}} \right) + 2}, \quad (11.10)$$

that shows how the dynamics deterministically drives the conditional state of the atomic sample into a spin-squeezed state.

In order to estimate the value of the magnetic field B , we consider the information obtained via the continuous weak-measurement via the coupled light modes, yielding the photocurrent (11.6), and the information obtained via a final “strong” measurement on the atomic ensemble, that, as explained just above, is eventually driven into a spin-squeezed state. In the next section we will show how one can quantify the performances of this measurement protocol exploiting tools of local quantum estimation theory.

11.3 Quantum Magnetometry via Time-Continuous Measurements

In this section we will assess the performances of the measurement protocol described in the previous section. In order to do this, we start by reviewing some basic notions on quantum estimation theory.

In a classical estimation problem, we want to infer the value of a parameter θ from a set of M measurement outcomes $\chi = \{x_1, x_2, \dots, x_M\}$ and their conditional distribution $p(x|\theta)$. In order to do so, we define an estimator $\hat{\theta}(\chi)$, that is, a function from the measurement outcomes to the possible values of θ . An estimator is called asymptotically unbiased when, in the limit of large M , its average is equal to the true value of θ , i.e. $\int d\chi p(\chi|\lambda) \hat{\theta}(\chi) = \theta$, where $p(\chi|\lambda) = \prod_{j=1}^M p(x_j|\theta)$.

The Cramér-Rao theorem gives a lower bound for the variance of any unbiased estimator:

$$\delta\theta^2 \geq \frac{1}{M\mathcal{F}[p(x|\theta)]}, \quad (11.11)$$

where

$$\mathcal{F}[p(x|\theta)] = \int dx p(x|\lambda)(\partial_\theta \log p(x|\lambda))^2 \quad (11.12)$$

is called Fisher information (FI).

In a quantum measurement setting, the conditional probability distribution reads $p(x|\theta) = \text{Tr}[\varrho_\theta \Pi_x]$, where ϱ_θ is the state of the quantum system, depending on the parameter θ , and Π_x is a POVM. One proves that the FI corresponding to any POVM is upper bounded

$$\mathcal{F}[p(x|\theta)] \leq \mathcal{Q}[\varrho_\theta],$$

where

$$\mathcal{Q}[\varrho_\theta] = 8 \lim_{d\theta \rightarrow 0} \frac{1 - F[\varrho_\theta, \varrho_{\theta+d\theta}]}{d\theta^2}$$

is the quantum Fisher information (QFI) expressed in terms of the fidelity between quantum states $F[\varrho, \sigma] = \text{Tr}[\sqrt{\sqrt{\sigma}\varrho\sqrt{\sigma}}]$ [50–52]. Notice that the QFI depends only on the quantum state ϱ_θ : It thus gives the ultimate bound on the precision of the estimation of θ allowed by quantum mechanics. Moreover, it can be shown that, when estimating a single parameter, the can always be achieved, i.e. one can find a (projective) POVM for which the classical FI equals the QFI.

In the quantum magnetometry protocol outlined in the previous section, one can identify two *sources* of information that can be exploited in order to estimate the magnetic field B . On one hand there is the homodyne photocurrent \mathbf{y}_t obtained via the weak measurement on the coupled light; on the other hand a final “strong” measurement can be performed, at the end of the dynamics, on the conditional state ϱ_c that evolves according to the stochastic master equation (11.6) and thus depends on the particular trajectory explored by the system due to the measurement. The information obtained via the photocurrent is quantified by the classical FI $\mathcal{F}[p(\mathbf{y}_t)]$ that can be calculated as described in [35] or, for continuous-variable Gaussian systems, as in [53]. The usefulness of the final strong measurement is similarly quantitatively expressed via the quantum Fisher information $\mathcal{Q}[\varrho_c]$. In [43] we proved that these two quantities can be combined to form a new quantum Cramér-Rao bound for time-continuous quantum magnetometry

$$\delta B^2 \geq \frac{1}{M (\mathcal{F}[p(\mathbf{y}_t)] + \mathbb{E}_{p(\mathbf{y}_t)} [\mathcal{Q}[\varrho^{(c)}])}]. \quad (11.13)$$

We can thus define an effective QFI that ultimately quantifies the estimation performances of our protocol

$$\tilde{\mathcal{Q}} = \mathcal{F}[p(\mathbf{y}_t)] + \mathbb{E}_{p(\mathbf{y}_t)} [\mathcal{Q}[\varrho^{(c)}]], \quad (11.14)$$

as the sum between the classical Fisher information $\mathcal{F}[p(\mathbf{y}_t)]$ and the average over all the possible trajectories of the conditional states QFI $\mathcal{Q}[\varrho^{(c)}]$.

In the large spin limit $J \gg 1$, both quantities can be analytically evaluated. In particular, following [53] one has

$$\begin{aligned} \mathcal{F}[p(\mathbf{y}_t)] &= 2\eta\kappa J e^{-\kappa t/2} \mathbb{E}_{p(\mathbf{y}_t)} \left[\left(\partial_B \langle \hat{P}(t) \rangle_c \right)^2 \right] \\ &= \frac{64\gamma^2 \eta J^2 e^{-\kappa t} \left(e^{\frac{\kappa t}{4}} - 1 \right)^3}{9\kappa^2 \left[(4\eta J + 1)e^{\frac{\kappa t}{2}} - 4\eta J \right]} \\ &\quad \cdot [-4\eta J - 12\eta J e^{\frac{\kappa t}{4}} + 3(4\eta J + 3)e^{\frac{\kappa t}{2}} + (4\eta J + 3)e^{\frac{3\kappa t}{4}}], \end{aligned} \quad (11.15)$$

while the conditional QFI can be evaluated as described in [54], obtaining

$$\begin{aligned} \mathcal{Q}[\varrho^{(c)}] &= \frac{\left(\partial_B \langle \hat{P}(t) \rangle_c \right)^2}{\text{Var}_c[\hat{P}(t)]} \\ &= \frac{32\gamma^2 J \left(12\eta J - 4\eta J e^{-\frac{\kappa t}{2}} + \frac{-8\eta J - 3}{\sqrt{e^{-\frac{\kappa t}{2}}}} + 3 \right)^2}{9\kappa^2 \left[(4\eta J + 1)e^{\frac{\kappa t}{2}} - 4\eta J \right]}. \end{aligned} \quad (11.16)$$

Since the conditional QFI does not depend on the particular trajectory defining the conditional state ϱ_c , the stochastic average present in (11.13) is not necessary, i.e. $\mathbb{E}_{p(\mathbf{y}_t)} [\mathcal{Q}[\varrho^{(c)}]] = \mathcal{Q}[\varrho^{(c)}]$, and the effective QFI can be readily evaluated.

By further manipulating the result, one obtains the following simple analytical formula

$$\tilde{\mathcal{Q}} = K_1 J + \eta K_2 J^2 \quad (11.17)$$

where

$$K_1 = 32 \frac{\gamma^2}{\kappa^2} \left(1 - e^{-\kappa t/4} \right)^2, \quad (11.18)$$

$$K_2 = 64 \frac{\gamma^2}{\kappa^2} \left(1 - \frac{8}{3} e^{-\kappa t/4} + 2e^{-\kappa t/2} - \frac{1}{3} e^{-\kappa t} \right). \quad (11.19)$$

In the following section we will discuss in more detail the significance of this formula, and the role played by the different experimental parameters describing the estimation protocol.

We quickly point out that it exists a more fundamental bound [36], optimized over all possible measurements on the environmental degrees of freedom. In [43] we

also prove that the effective QFI (11.17) is optimal in this sense, as it saturates the fundamental bound of [36] for $\eta = 1$.

11.3.1 The Role of Monitoring Time and Measurement Efficiency

The formula reported in (11.17) is particularly illuminating and in particular it allows to discuss in more detail the role played by the two parameters describing the protocol: the (dimensionless) monitoring time κt and the monitoring efficiency η . We start by observing that the effective QFI is the sum of two terms presenting respectively a SQL scaling proportional to J , and a Heisenberg scaling proportional to J^2 . The latter one depends linearly on the monitoring efficiency η and correctly goes to zero for $\eta \rightarrow 0$, i.e. when no monitoring is performed and the SQL-scaling is the expected result for an initial spin-coherent state. To better analyze the behaviour of the effective QFI, we plot it in Fig. 11.1 as function of J for fixed monitoring time κt and efficiency η . In the figures we indeed observe a transition from SQL-like scaling to Heisenberg scaling depending on the values of J , η and κt . In particular, they show quantum enhancement for $J \gg 1/\kappa t$. This observation can be rigorously studied by looking

Fig. 11.1 Effective QFI \tilde{Q} as a function of J , with $\gamma/\kappa = 1 \text{ G}^{-1}$. We show a dashed line at the top $\propto J^2$ and a dotted line at the bottom $\propto J$ as visual guides. (Top panel): fixed unit efficiency $\eta = 1$; the curves represent increasing values of κt from top to bottom (see the legend). (Bottom panel): fixed monitoring time $\kappa t = 10^{-2}$; the curves represent increasing values of η from top to bottom (see the legend)

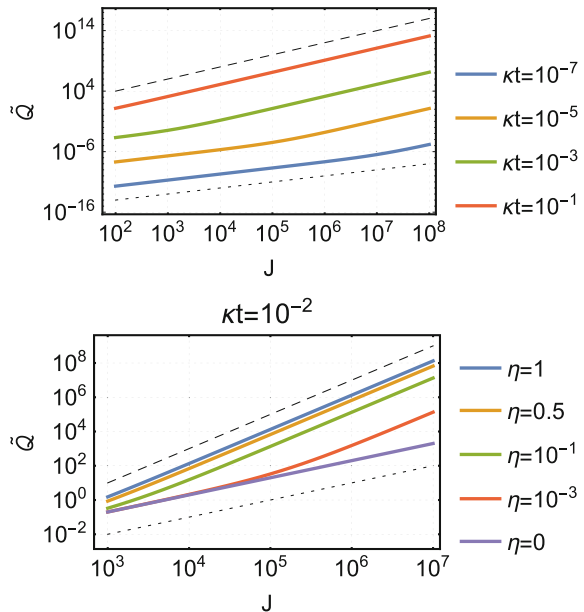
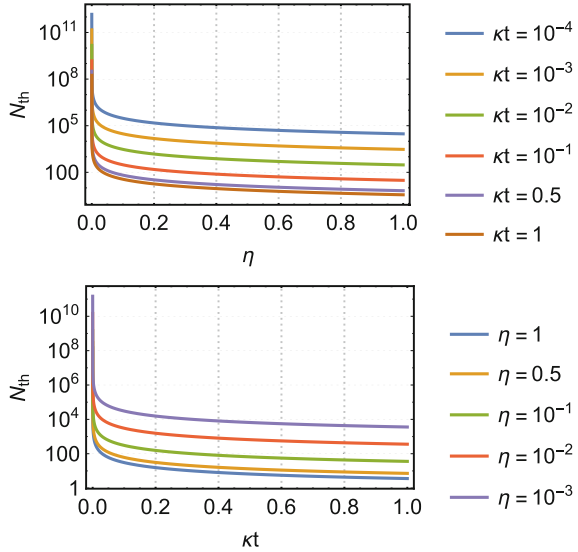


Fig. 11.2 The minimum number N_{th} of atoms in the ensemble needed to observe Heisenberg scaling for the effective QFI \tilde{Q} . In the left panel we plot N_{th} as a function of η for different values of κt , there is an inverse proportionality relationship. In the right panel we plot the same quantity as a function of κt for different values of η , in this case the relationship is only approximately inversely proportional



more carefully at (11.17), (11.18) and (11.19). One can in fact define the condition to observe Heisenberg scaling as

$$\text{Heisenberg scaling} \iff \eta K_2 J > K_1 \tag{11.20}$$

i.e. when the quadratic contribution is greater than the linear one. This implies that we have a threshold value for the total spin,

$$J_{\text{th}} = \frac{K_1}{\eta K_2}, \tag{11.21}$$

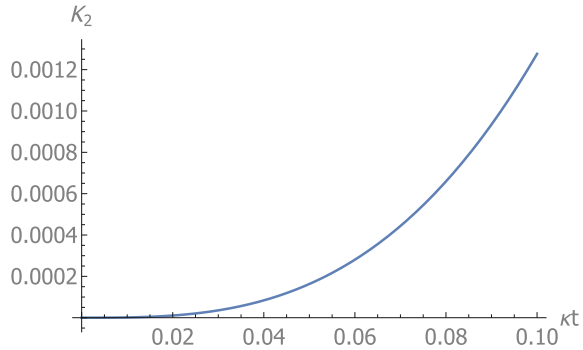
and obviously, in terms of the total number of atoms, $N_{\text{th}} = 2J_{\text{th}}$. If we plug in the coefficients from (11.18) and (11.19), we get this simple expression

$$N_{\text{th}} = \frac{1}{\eta \left(1 - \frac{1}{3}e^{-\kappa t/2} - \frac{2}{3}e^{-\kappa t/4} \right)}, \tag{11.22}$$

which can be approximated as $N_{\text{th}} \approx \frac{3}{\eta \kappa t}$ when $\kappa t \ll 1$. The expected inverse relationship between the threshold number of atoms and both the monitoring time κt and the efficiency η , can be explicitly observed in Fig. 11.2, where we plot N_{th} as a function of both experimental parameters characterizing the protocol.

Once the Heisenberg scaling is recovered, it is also interesting to study the behaviour of the asymptotic constant factor K_2 , corresponding to the Heisenberg-like term in the effective QFI. We plot it in Fig. 11.3 as a function of monitoring

Fig. 11.3 Asymptotic constant factor K_2 for the Heisenberg-like term in the effective QFI \tilde{Q} as a function of monitoring time κt (for fixed $\gamma/\kappa = 1$)



time κt : we observe, as expected a monotonous behaviour, that for $\kappa t \ll 1$ can be approximated by the formula

$$K_2 \approx \frac{4}{3} \frac{\gamma^2}{\kappa^2} (\kappa t)^3,$$

where one can observe the κt^3 behaviour that was already highlighted in [43].

11.4 Conclusions and Perspectives

We have analytically evaluated the ultimate precision achievable for the estimation of a constant weak magnetic field acting on an atomic ensemble that is continuously monitored via homodyne detection of a coupled electromagnetic mode. We have shown that under some conditions, one obtains a Heisenberg scaling, even starting from an initial “classical” spin-coherent state. Remarkably, we have been able to obtain an analytical condition that the experimental parameters must satisfy in order to observe the desired quantum enhancement. In particular we have shown how, for finite monitoring time and monitoring efficiency, there is a threshold on the total number of atoms above which one observes a transition from SQL to Heisenberg scaling. Time-continuous monitoring is a promising tool for quantum metrology. We leave to future investigations the study of its usefulness in the presence of more detrimental noisy dynamics (such as the one corresponding to independent noise on each atom) and in the multi-parameter scenario.

References

1. V. Giovannetti, S. Lloyd, L. Maccone, Nat. Photonics **5**, 222 (2011). <https://doi.org/10.1038/nphoton.2011.35>. <http://www.nature.com/doi/10.1038/nphoton.2011.35>

2. R. Demkowicz-Dobrzanski, M. Jarzyna, J. Kolodynski, *Prog. Opt.* **60**, 345 (2015). <https://doi.org/10.1016/bs.po.2015.02.003>. <http://arxiv.org/abs/1405.7703>linkinghub.elsevier.com/retrieve/pii/S0079663815000049
3. W. Wasilewski, K. Jensen, H. Krauter, J.J. Renema, M.V. Balabas, E.S. Polzik, *Phys. Rev. Lett.* **104**(13), 133601 (2010). <https://doi.org/10.1103/PhysRevLett.104.133601>. <https://link.aps.org/doi/10.1103/PhysRevLett.104.133601>
4. M. Koschorreck, M. Napolitano, B. Dubost, M.W. Mitchell, *Phys. Rev. Lett.* **104**(9), 093602 (2010). <https://doi.org/10.1103/PhysRevLett.104.093602>. <https://link.aps.org/doi/10.1103/PhysRevLett.104.093602>
5. R.J. Sewell, M. Koschorreck, M. Napolitano, B. Dubost, N. Behbood, M.W. Mitchell, *Phys. Rev. Lett.* **109**(25), 253605 (2012). <https://doi.org/10.1103/PhysRevLett.109.253605>. <https://link.aps.org/doi/10.1103/PhysRevLett.109.253605>
6. C.F. Ockeloen, R. Schmied, M.F. Riedel, P. Treutlein, *Phys. Rev. Lett.* **111**(14), 143001 (2013). <https://doi.org/10.1103/PhysRevLett.111.143001>. <https://link.aps.org/doi/10.1103/PhysRevLett.111.143001>
7. D. Sheng, S. Li, N. Dural, M.V. Romalis, *Phys. Rev. Lett.* **110**(16), 160802 (2013). <https://doi.org/10.1103/PhysRevLett.110.160802>. <https://link.aps.org/doi/10.1103/PhysRevLett.110.160802>
8. V.G. Lucivero, P. Anielski, W. Gawlik, M.W. Mitchell, *Rev. Sci. Instrum.* **85**(11), 113108 (2014). <https://doi.org/10.1063/1.4901588>. <http://arxiv.org/abs/1403.7796>, <http://dx.doi.org/10.1063/1.4901588>, <http://aip.scitation.org/doi/10.1063/1.4901588>
9. W. Muessel, H. Strobel, D. Linnemann, D.B. Hume, M.K. Oberthaler, *Phys. Rev. Lett.* **113**(10), 103004 (2014). <https://doi.org/10.1103/PhysRevLett.113.103004>. <https://link.aps.org/doi/10.1103/PhysRevLett.113.103004>
10. D.J. Wineland, J.J. Bollinger, W.M. Itano, F.L. Moore, D.J. Heinzen, *Phys. Rev. A* **46**(11), R6797 (1992). <https://doi.org/10.1103/PhysRevA.46.R6797>. <http://link.aps.org/doi/10.1103/PhysRevA.46.R6797>
11. J.J. Bollinger, W. Itano, D.J. Wineland, D.J. Heinzen, *Phys. Rev. A* **54**(6), R4649 (1996). <https://doi.org/10.1103/PhysRevA.54.R4649>
12. S.F. Huelga, C. Macchiavello, T. Pellizzari, A.K. Ekert, M.B. Plenio, J.I. Cirac, *Phys. Rev. Lett.* **79**(20), 3865 (1997). <https://doi.org/10.1103/PhysRevLett.79.3865>. <https://link.aps.org/doi/10.1103/PhysRevLett.79.3865>
13. B.M. Escher, R.L. de Matos Filho, L. Davidovich, *Nat. Phys.* **7**(5), 406 (2011). <https://doi.org/10.1038/nphys1958>. <http://dx.doi.org/10.1038/nphys1958>, <http://www.nature.com/doi/10.1038/nphys1958>
14. R. Demkowicz-Dobrzanski, J. Kolodynski, M. Guta, *Nat. Commun.* **3**, 1063 (2012). <https://doi.org/10.1038/ncomms2067>. <http://www.nature.com/doi/10.1038/ncomms2067>
15. Y. Matsuzaki, S.C. Benjamin, J. Fitzsimons, *Phys. Rev. A* **84**(1), 012103 (2011). <https://doi.org/10.1103/PhysRevA.84.012103>. <https://link.aps.org/doi/10.1103/PhysRevA.84.012103>
16. A.W. Chin, S.F. Huelga, M.B. Plenio, *Phys. Rev. Lett.* **109**(23), 233601 (2012). <https://doi.org/10.1103/PhysRevLett.109.233601>. <https://link.aps.org/doi/10.1103/PhysRevLett.109.233601>
17. R. Chaves, J.B. Brask, M. Markiewicz, J. Kolodynski, A. Acin, *Phys. Rev. Lett.* **111**(12), 120401 (2013). <https://doi.org/10.1103/PhysRevLett.111.120401>. <https://link.aps.org/doi/10.1103/PhysRevLett.111.120401>
18. J.B. Brask, R. Chaves, J. Kolodynski, *Phys. Rev. X* **5**(3), 031010 (2015). <https://doi.org/10.1103/PhysRevX.5.031010>. <http://arxiv.org/abs/1411.0716>, <http://link.aps.org/doi/10.1103/PhysRevX.5.031010>
19. A. Smirne, J. Kolodynski, S.F. Huelga, R. Demkowicz-Dobrzanski, *Phys. Rev. Lett.* **116**(12), 120801 (2016). <https://doi.org/10.1103/PhysRevLett.116.120801>. <http://link.aps.org/doi/10.1103/PhysRevLett.116.120801>
20. E.M. Kessler, I. Lovchinsky, A.O. Sushkov, M.D. Lukin, *Phys. Rev. Lett.* **112**(15), 150802 (2014). <https://doi.org/10.1103/PhysRevLett.112.150802>. <https://link.aps.org/doi/10.1103/PhysRevLett.112.150802>

21. W. Dur, M. Skotiniotis, F. Frowis, B. Kraus, Phys. Rev. Lett. **112**(8), 080801 (2014). <https://doi.org/10.1103/PhysRevLett.112.080801>. <http://link.aps.org/doi/10.1103/PhysRevLett.112.080801>
22. G. Arrad, Y. Vinkler, D. Aharonov, A. Retzker, Phys. Rev. Lett. **112**(15), 150801 (2014). <https://doi.org/10.1103/PhysRevLett.112.150801>. <https://link.aps.org/doi/10.1103/PhysRevLett.112.150801>
23. P. Sekatski, M. Skotiniotis, J. Kolodynski, W. Dur, Quantum. <https://doi.org/10.22331/q-2017-09-06-27>. <https://doi.org/10.22331/q-2017-09-06-27>
24. T. Unden, P. Balasubramanian, D. Louzon, Y. Vinkler, M.B. Plenio, M. Markham, D. Twitchen, A. Stacey, I. Lovchinsky, A.O. Sushkov, M.D. Lukin, A. Retzker, B. Naydenov, L.P. McGuinness, F. Jelezko, Phys. Rev. Lett. **116**, 230502 (2016). <https://doi.org/10.1103/PhysRevLett.116.230502>. <https://link.aps.org/doi/10.1103/PhysRevLett.116.230502>
25. S. Zhou, M. Zhang, J. Preskill, L. Jiang, ArXiv e-prints (2017)
26. R. Demkowicz-Dobrzanski, J. Czajkowski, P. Sekatski, ArXiv e-prints (2017)
27. H.M. Wiseman, G.J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, New York, 2010)
28. K. Jacobs, D.A. Steck, Contemp. Phys. **47**(5), 279 (2006). <https://doi.org/10.1080/00107510601101934>. <http://www.tandfonline.com/doi/abs/10.1080/00107510601101934>
29. J.M. Geremia, J.K. Stockton, A.C. Doherty, H. Mabuchi, Phys. Rev. Lett. **91**(25), 250801 (2003). <https://doi.org/10.1103/PhysRevLett.91.250801>. <http://link.aps.org/doi/10.1103/PhysRevLett.91.250801>
30. J.K. Stockton, J.M. Geremia, A.C. Doherty, H. Mabuchi, Phys. Rev. A **69**(3), 032109 (2004). <https://doi.org/10.1103/PhysRevA.69.032109>. <https://link.aps.org/doi/10.1103/PhysRevA.69.032109>
31. M. Auzinsh, D. Budker, D.F. Kimball, S.M. Rochester, J.E. Stalnaker, A.O. Sushkov, V.V. Yashchuk, Phys. Rev. Lett. **93**(17), 173002 (2004). <https://doi.org/10.1103/PhysRevLett.93.173002>. <https://link.aps.org/doi/10.1103/PhysRevLett.93.173002>
32. K. Molmer, L.B. Madsen, Phys. Rev. A **70**(5), 052102 (2004). <https://doi.org/10.1103/PhysRevA.70.052102>. <https://link.aps.org/doi/10.1103/PhysRevA.70.052102>
33. L.B. Madsen, K. Molmer, Phys. Rev. A **70**(5), 052324 (2004). <https://doi.org/10.1103/PhysRevA.70.052324>. <https://link.aps.org/doi/10.1103/PhysRevA.70.052324>
34. B.A. Chase, J.M. Geremia, Phys. Rev. A **79**(2), 022314 (2009). <https://doi.org/10.1103/PhysRevA.79.022314>. <https://link.aps.org/doi/10.1103/PhysRevA.79.022314>
35. S. Gammelmark, K. Molmer, Phys. Rev. A **87**(3), 032115 (2013). <https://doi.org/10.1103/PhysRevA.87.032115>. <http://link.aps.org/doi/10.1103/PhysRevA.87.032115>
36. S. Gammelmark, K. Molmer, Phys. Rev. Lett. **112**(17), 170401 (2014). <https://doi.org/10.1103/PhysRevLett.112.170401>. <http://link.aps.org/doi/10.1103/PhysRevLett.112.170401>
37. A.H. Kiilerich, K. Molmer, Phys. Rev. A **89**(5), 052110 (2014). <https://doi.org/10.1103/PhysRevA.89.052110>. <http://link.aps.org/doi/10.1103/PhysRevA.89.052110>
38. A.H. Kiilerich, K. Molmer, Phys. Rev. A **94**(3), 032103 (2016). <https://doi.org/10.1103/PhysRevA.94.032103>. <http://link.aps.org/doi/10.1103/PhysRevA.94.032103>
39. C. Catana, M. Guta, Phys. Rev. A **90**(1), 012330 (2014). <https://doi.org/10.1103/PhysRevA.90.012330>. <https://link.aps.org/doi/10.1103/PhysRevA.90.012330>
40. T. Gefen, D.A. Herrera-Martí, A. Retzker, Phys. Rev. A **93**(3), 032133 (2016). <https://doi.org/10.1103/PhysRevA.93.032133>. <http://link.aps.org/doi/10.1103/PhysRevA.93.032133>
41. M.B. Plenio, S.F. Huelga, Phys. Rev. A **93**(3), 032123 (2016). <https://doi.org/10.1103/PhysRevA.93.032123>. <https://link.aps.org/doi/10.1103/PhysRevA.93.032123>
42. L. Cortez, A. Chantasri, L.P. Garcia-Pintos, J. Dressel, A.N. Jordan, Phys. Rev. A **95**(1), 012314 (2017). <https://doi.org/10.1103/PhysRevA.95.012314>. <http://arxiv.org/abs/1606.01407>, <http://link.aps.org/doi/10.1103/PhysRevA.95.012314>
43. F. Albarelli, M.A.C. Rossi, M.G.A. Paris, M.G. Genoni, (2017). <http://arxiv.org/abs/1706.00485>
44. L.K. Thomsen, S. Mancini, H.M. Wiseman, Phys. Rev. A **65**(6), 061801 (2002). <https://doi.org/10.1103/PhysRevA.65.061801>. <https://link.aps.org/doi/10.1103/PhysRevA.65.061801>

45. D. Plankensteiner, J. Schachenmayer, H. Ritsch, C. Genes, J. Phys. B **49**, 245501 (2016). <https://doi.org/10.1088/0953-4075/49/24/245501>. <http://arxiv.org/abs/1605.00874>stacks.iop.org/0953-4075/49/i=24/a=245501?key=crossref.8d42d18e21fd8245d1b3c7a202188cc1
46. E.G. Dalla Torre, J. Otterbach, E. Demler, V. Vuletic, M.D. Lukin, Phys. Rev. Lett. **110**, 120402 (2013). <https://doi.org/10.1103/PhysRevLett.110.120402>. <https://link.aps.org/doi/10.1103/PhysRevLett.110.120402>
47. U. Dörner, New J. Phys. **14**(4), 043011 (2012). <https://doi.org/10.1088/1367-2630/14/4/043011>. <http://stacks.iop.org/1367-2630/14/i=4/a=043011?key=crossref.50d99a3afee8de2c2f336d63e774ebcf>
48. M.G. Genoni, L. Lami, A. Serafini, Contemp. Phys. **57**(3), 331 (2016). <https://doi.org/10.1080/00107514.2015.1125624>. <http://www.tandfonline.com/doi/full/10.1080/00107514.2015.1125624>
49. H.M. Wiseman, A.C. Doherty, Phys. Rev. Lett. **94**(7), 070405 (2005). <https://doi.org/10.1103/PhysRevLett.94.070405>. <https://link.aps.org/doi/10.1103/PhysRevLett.94.070405>
50. C.W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976)
51. S.L. Braunstein, C.M. Caves, Phys. Rev. Lett. **72**(22), 3439 (1994). <https://doi.org/10.1103/PhysRevLett.72.3439>. <https://link.aps.org/doi/10.1103/PhysRevLett.72.3439>
52. M.G.A. Paris, Int. J. Quant. Inf. **07**(supp01), 125 (2009). <https://doi.org/10.1142/S0219749909004839>, <http://www.worldscientific.com/doi/abs/10.1142/S0219749909004839>
53. M.G. Genoni, Phys. Rev. A **95**(1), 012116 (2017). <https://doi.org/10.1103/PhysRevA.95.012116>. <https://link.aps.org/doi/10.1103/PhysRevA.95.012116>
54. O. Pinel, P. Jian, N. Treps, C. Fabre, D. Braun, Phys. Rev. A **88**(4), 040102 (2013). <https://doi.org/10.1103/PhysRevA.88.040102>. <https://link.aps.org/doi/10.1103/PhysRevA.88.040102>

Chapter 12

Recent Developments and Perspectives in Nuclear Structure by γ and Particle Spectroscopy



S. Leoni, A. Bracco, F. Camera, A. Guglielmetti, G. Benzoni, B. Blasi, C. Boiano, Pier Francesco Bortignon, S. Bottoni, S. Brambilla, S. Capra, G. Colò, F. C. L. Crespi, A. Mentana, B. Million, A. Pullia, S. Riboldi, E. Vigezzi and O. Wieland

Abstract Advances in nuclear structures studies of neutron-rich systems are presented, focusing on experimental activities based on γ and particle spectroscopy—ideal probes of the multifaceted nature of the atomic nucleus. Emphasis is given to recent highlights obtained in world’s leading laboratories, in Europe and Japan, by large international collaborations in which the Nuclear Physics group of Milano University has played a key role. In particular, experiments are described taking advantage of both stable and radioactive ion beams, as well as intense neutron beams. They require the use of the most advanced detection setups, such as the European array AGATA, based on γ -ray tracking techniques, and also arrays of large size scintillators. Measurements of cross sections of thermonuclear fusion reactions at astrophysically relevant energies, performed by the LUNA collaboration, are also briefly discussed. Finally, research and developments of detectors based on novel scintillators materials are presented, being a significant part of the activity carried out at the Milano Physics Department.

Pier Francesco Bortignon has deceased.

S. Leoni (✉) · A. Bracco · F. Camera · A. Guglielmetti · P. F. Bortignon · S. Bottoni · S. Capra · G. Colò · F. C. L. Crespi · A. Mentana · A. Pullia · S. Riboldi
Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy
e-mail: silvia.leoni@mi.infn.it

S. Leoni · A. Bracco · F. Camera · A. Guglielmetti · G. Benzoni · B. Blasi · C. Boiano · P. F. Bortignon · S. Bottoni · S. Brambilla · S. Capra · G. Colò · F. C. L. Crespi · A. Mentana · B. Million · A. Pullia · S. Riboldi · E. Vigezzi · O. Wieland
INFN, Sezione di Milano, Via Celoria 16, 20133 Milan, Italy

© Springer Nature Switzerland AG 2018
P. F. Bortignon et al. (eds.), *Toward a Science Campus in Milan*,
https://doi.org/10.1007/978-3-030-01629-6_12

12.1 Challenges in Nuclear Structures

The atomic nucleus, consisting of densely packed protons and neutrons, is a quantum system in which three fundamental interactions—strong, electromagnetic and weak—can be studied. Despite a great amount of experimental data collected for more than 3000 nuclei observed in laboratories on Earth—out of more than 7000 expected to exist in the Universe (see Fig. 12.1)—a unified and satisfactory theoretical description of the properties of the atomic nucleus is still missing. This is a consequence of the complex nature of the nucleon-nucleon interaction and of the interplay of several degrees of freedom, in particular (i) angular momentum (up to the fission limit), (ii) temperature (i.e., internal excitation energy, up to and above particle-binding energy), and (iii) isospin (i.e. the neutron to proton ratio, N/Z , indicating the location with respect to the stability valley). This leads to the appearance of a rich variety of phenomena, such as different modes of excitations (involving few nucleons and/or collective responses) or coexistence of different nuclear shapes and deformations, posing a challenge to the most sophisticated theoretical descriptions.

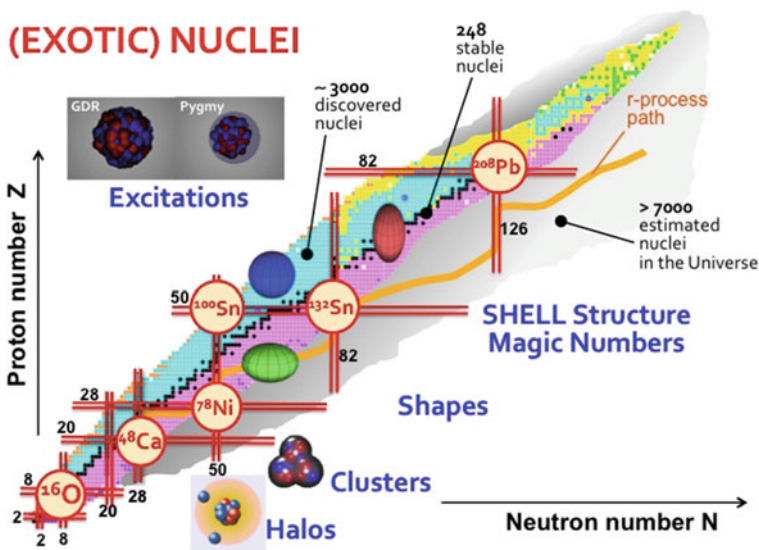


Fig. 12.1 Pictorial view of the nuclear chart, indicating the stability valley (in black) and known unstable nuclei (β^+ , β^- and α emitters in light blue, pink and yellow, respectively). The grey area shows the region of still unknown, unstable “exotic” nuclei expected to exist, for a total of more than 7000 atomic nuclei in the Universe. Proton and neutron shell closures are indicated by solid red lines, together with doubly magic systems, such as ^{16}O , ^{78}Ni , ^{48}Ca , ^{100}Sn , ^{132}Sn and ^{208}Pb . Examples of specific structures, like halos and clusters, spherical, oblate and prolate deformed shapes and collective Giant and Pygmy dipole resonance excitations are given by cartoons. The orange line refer to the rapid neutron capture path (i.e., the r-process path), which is responsible for the formation of elements heavier than ^{56}Fe [1]

In this brief review, selected examples are discussed among the most significant and recent experimental investigations carried out by the Nuclear Physics group of Milano University, with close collaboration between theory and experiment. In particular, studies of coexistence of shapes, couplings between single particle and collective degree excitations (phonons) and resonances are presented, often in connection with exotic, n-rich nuclei (away from stability), which are at the forefront of the experimental and theoretical investigations and strongly influence reaction rates in stellar environment (cfr. Fig. 12.1) [1]. The research activities are conducted within large international collaborations, in world's leading laboratories, in Europe (e.g., LNL (Padova), GSI (Darmstadt), ILL (Grenoble), ISOLDE (CERN), Bucharest...), and Japan (RIKEN and OSAKA). Reactions with both stable and radioactive ions beam (from low to relativistic energies) and intense neutron beams are exploited. Very complex detection setup are used, most often combining state-of-the-art arrays based on high purity Ge crystals (such as AGATA, the Advanced GAMMA Tracking Array [2]) and large size scintillator detectors of novel type, being γ and particle spectroscopy the main investigation techniques.

Nuclear astrophysics activities, on-going at the underground Gran Sasso Laboratory, are also briefly discussed. Finally, the last section is devoted to technical research studies related to novel scintillator materials, being an important part of the research program of the Milano Nuclear Physics group.

12.1.1 Shape Coexistence and Shape Isomerism

The shape is one of the most fundamental properties of the nucleus: while spherical shapes are most natural in the vicinities of doubly shell closures, in the regions lying away from doubly magic nuclei different nuclear shapes—with dominance of quadrupole symmetric forms—are competing and may coexist in the same nucleus, at low excitation energy and low spins [3, 4]. Further, superdeformed and spherical states are found to coexist at high energy and high spin in open-shell nuclei [5–7].

In even-even systems, most striking fingerprints of shape coexistence are low lying 0^+ excited states associated with deformations different from the ground state. Taking advantage of various experimental probes (e.g., γ -ray and conversion-electron measurements and laser spectroscopy), they have been discovered in different regions of the nuclear chart: from the rather “light” Si/Mg and Ni/Zn/Ge to the medium mass Kr/Rb/Sr/Zr isotopes, to the much heavier Po/Pb/Hg and Rn/Ra systems. A very peculiar manifestation of the coexistence of shapes are shape isomers: they may arise when the nuclear potential energy surface (PES) in the deformation space has minima associated with different shapes and when these minima are separated by a high barrier [8]. So far, shape isomers, discovered in the 60s, were known to occur, clearly, in the heavy actinides only: there, in two cases— ^{236}U and ^{238}U —very retarded E2 γ branches from excited 0^+ states were observed [9–11], although, since the 80s, mean-field models predicted the existence of shape isomers also in lighter systems, pointing to ^{66}Ni and ^{68}Ni as the lightest candidates [12–14]. Indeed,

extended experimental campaigns performed at ISOLDE/CERN, MSU and RIKEN have confirmed coexistence of spherical, oblate and prolate shapes along the Ni isotopic chain, although no traces of shape isomers were found in ^{68}Ni and heavier systems [15–17].

Recently, fully microscopic, state-of-the-art shell-model calculations, based on ingenious (Monte Carlo) computational schemes and the use of very powerful super-computing systems (the Japanese K-computer with 1 million parallel processors), predicted shape coexistence in neutron-rich Ni isotopes and indicated ^{66}Ni as the best case for the existence of shape isomerism [18]. This prediction was confirmed by a measurement of the lifetimes of 0^+ excitations in ^{66}Ni , performed at the Bucharest Tandem Laboratory. By employing the two-neutron transfer reaction $^{64}\text{Ni}(^{18}\text{O}, ^{16}\text{O})^{66}\text{Ni}$, at sub-barrier energy of 39 MeV, the γ -decay from all three lowest-excited 0^+ states in ^{66}Ni was observed with the ROSPHERE HPGe array [19].

The population pattern clearly indicated that 0_4^+ corresponds to the prolate deformed 0^+ excitation, predicted by theory, and its lifetime, measured with a plunger device, gave a $B(E2)$ value of 0.2 W.u. (“or single-particle units”), for the 0_4^+ to 2_1^+ decay. This result makes ^{66}Ni a unique nuclear system (apart from the much heavier ^{236}U and ^{238}U) in which a retarded E2 γ transition from a 0^+ deformed state to a spherical configuration is observed, pointing to a shape isomer-like behavior. As a consequence, it becomes clear that shape isomerism remarkably appears also in significantly lighter systems than the actinides. Moreover, the proper theory prediction represents a significant step forward in understanding the microscopic origin

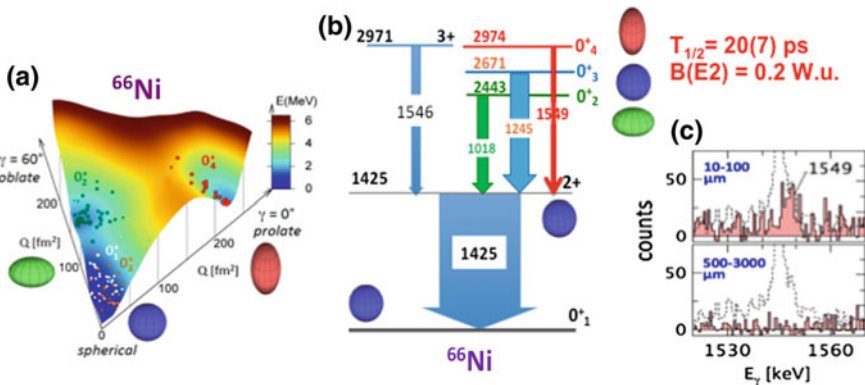


Fig. 12.2 Panel **a** Potential Energy Surface (PES) for the lowest 0^+ states of ^{66}Ni , showing minima associated with different shapes: spherical, oblate and prolate. Circles represent shapes in the Monte Carlo Shell Model basis vectors [18]. Panel **b** Partial level scheme of ^{66}Ni , relevant for the γ decay of 0^+ excitations (lifetime and $B(E2)$ of the shape-isomer-like 0_4^+ state are indicated in red). Panel **c** portions of γ spectra of ^{66}Ni , showing the photon decay from the shape-isomer-like state. Top: γ emission in the time range 5–15 ps; Bottom: γ emission in the time range 75–450 ps. Dashed lines: Time integrated γ -ray spectrum. (adapted from [19])

of nuclear deformation, within a pure shell model approach, a key issue in nuclear structure physics (Fig. 12.2).

12.1.2 *Couplings Between Nucleons and Core Excitations*

The interplay between single particle excitations and collective responses of the nucleus generates a multifaceted scenario of nuclear excitations. Ideal systems for their studies—in the simplest form—are nuclei made of one valence particle and a doubly magic core. In these cases, long range correlations, such as couplings between particle and excitations of the core (phonons in particular) are major sources of partial occupancies of nucleonic orbitals (as evidenced by knock out and transfer reactions) [20], they are doorways to the damping of resonance excitations [21] and have an impact on the Gamow-Teller strength function in the β -decay of magic systems [22].

Recently, extended γ -spectroscopy studies were performed for the one-valence neutron and proton nuclei ^{49}Ca and ^{133}Sb , respectively. In the first case, the nucleus of interest, ^{49}Ca , was populated by the heavy-ion deep inelastic reaction $^{48}\text{Ca} + ^{64}\text{Ni}$, at ~ 6 MeV/nucleon [23, 24]. The experiment was performed at Legnaro National Laboratory of INFN (Padova), using the large acceptance magnetic spectrometer PRISMA [25] combined with the high efficiency HPGe array CLARA. In the second case, the ^{133}Sb nucleus was produced by neutron induced fission of ^{235}U and ^{241}Pu targets, at Institut Laue-Langevin (ILL, Grenoble), where a reactor delivers the world brightest continuous neutron beam, available for research studies. The γ decay of ^{133}Sb was measured with a highly efficient HPGe array, coupled to fast LaBr_3 scintillator detectors [26].

In both experiments, the lifetimes of selected excited states was measured and the transition probability of the electromagnetic decay between excited states was deduced, providing clear indication of the nature of the states. In the case of ^{49}Ca , the $9/2^+$ state was found to arise from the coupling of the odd neutron to the 3^- octupole phonon of the ^{48}Ca core, while in ^{133}Sb a more complex picture was found. In this case, a fast change in the nature of the excitations, with increasing spin, was observed, pointing to couplings between the valence proton and excitations of the ^{132}Sn core, of both genuine phonon type and less collective character [27]. The hybrid nature of the excitations in ^{133}Sb is illustrated by the histograms of Fig. 12.3, giving the composition of the state configuration, sketched on the corresponding cartoons on top of the figure. Hybridisation phenomena of the type reported here are well known in all branches of physics, including condensed matter: there, electrons and plasmons in metals [28], for example, are the counterparts of the proton and collective excitation of the ^{132}Sn core, in the ^{133}Sb nucleus.

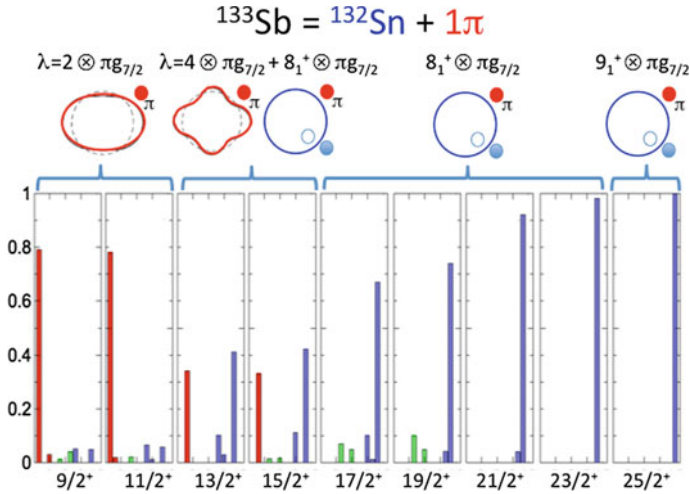


Fig. 12.3 Evolution of the yrast states composition in ^{133}Sb : the low spin states are dominated by the valence $g_{7/2}$ proton coupled to the 2^+ phonon of the ^{132}Sn core, highest spin excitations by the coupling to the neutron $h_{11/2}^- f_{7/2}$ non-collective core excitation [26]

12.1.3 Collective Excitations: Giant and Pygmy Resonances

One of the most extreme manifestations of collective excitations are Giant Resonances, which involve a large fraction of constituent nucleons [21, 29]. They provide information on bulk properties of nuclei and their measurements in exotic systems is technically very challenging. In neutron-rich systems, the neutron excess forms a skin, which is often assumed to oscillate outside the proton-neutron core. This results in a concentration of electric dipole E1 strength in the region around the particle binding energy (<10 MeV)—the Pygmy Dipole Resonance (PDR)—which strongly influences reaction rates in the astrophysical r-process (see Fig. 12.1). Few experiments are ongoing for exotic nuclei and above separation threshold, using advanced and complex setups at fragmentation facilities (e.g., GSI, RIKEN, ...). Figure 12.4a shows the first, pioneering measurement of the γ decay from pygmy states by Wieland et al. [30]. It was obtained by Coulomb excitation of ^{68}Ni at 600 MeV/nucleon on a Au target, using the RISING setup (comprising large volume HPGe crystals as well as BaF₂ scintillators) at the fragment separator of GSI. Evidence was found for a strength of about 5% of the energy weighed sum rule (EWSR) for the E1 isovector decay, significantly larger than in stable systems.

Contrary to exotic nuclei, stable neutron-rich systems, below threshold, have been extensively investigated by different probes—from photons, to protons and alphas, to heavy ions—revealing a quite complex nature of pygmy states: in the same energy region isoscalar and isovector states seem to co-exist, and the character of these excitations appears to be hybrid, with mixture of compressional and/or non-collective character [31]. Figure 12.4c–e show the results of a systematic investigation per-

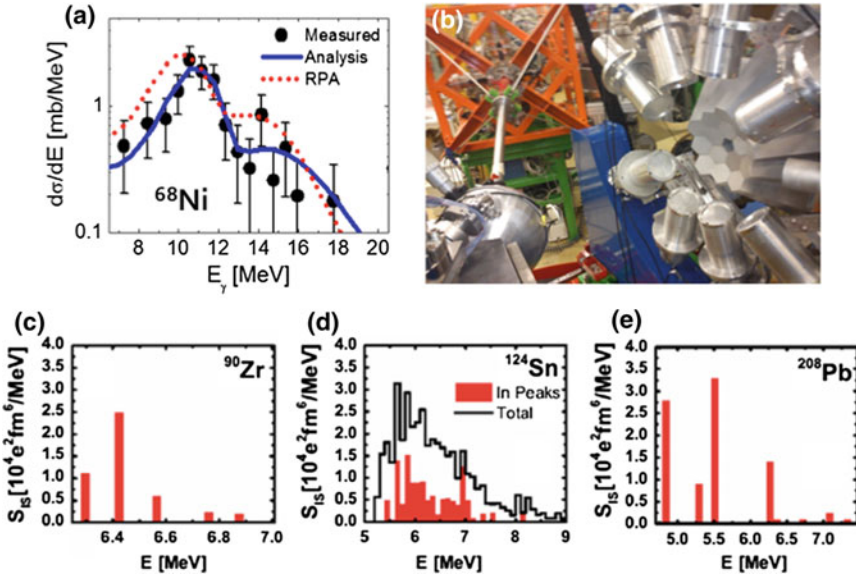


Fig. 12.4 Overview of selected experimental investigations of the γ decay of the Pygmy Dipole Resonance in neutron rich systems. Panel **a** Coulomb excitation of the exotic nucleus ^{68}Ni , measured at the GSI fragment separator at 600 MeV/nucleon [30], Panels **c–e** isoscalar strength distribution for the Pygmy excitations in ^{90}Zr , ^{124}Sn and ^{208}Pb stable nuclei, respectively, as measured at LNL by inelastic scattering of ^{17}O ions at 20 MeV/nucleon. Red bars refer to discrete transitions, while the black histogram gives, in the case of ^{124}Sn , the total strength (including the unresolved part (adapted from [31])). The used setup, the AGATA Demonstrator array coupled to large size LaBr₃:Ce scintillators, is shown in panel **(b)**

formed at Legnaro National Laboratory of INFN (Padova, Italy), using the (^{17}O , $^{17}\text{O}'$ γ) reaction at 20 MeV/nucleon [32–35] and a setup comprising the first implementation of the most advanced HPGe array—the Advanced Gamma Tracking Array (AGATA) [2]—and a number of LaBr₃:Ce, large volume, scintillator detectors (see panel **b**)). The features of the low-lying dipole strength will be better clarified in the future, moving along isotopic chains, towards more neutron rich systems, as well as the presence of pygmy states of other multipolarities, such as E2, for which indications already exist in ^{124}Sn [36]. Finally, experiments based on the detection of high resolution γ rays and particle decay at intense γ -beam facilities, such as ELI-NP [37], will allow us to get access, in stable systems, to the fine structure of the resonance response, over the entire excitation energy spectrum, shedding light on damping mechanisms.

12.1.4 β -Decay Studies of Exotic Nuclei

A useful tool to access information on nuclei far from stability is the study of the β decay of exotic nuclei. Gross properties of β -decay, such as the decay half-life and branching ratios, can be obtained already with a limited number of ions, offering the first insight in the structure of very exotic nuclei and necessary inputs to understand stellar nucleosynthesis processes (e.g., the r-process shown in Fig. 12.1).

Campaigns to study β -decay properties of n-rich nuclei are being performed in all facilities producing radioactive beams in Europe, USA and Japan. In a recent experiment performed at the ISOLDE Decay Station (IDS), located at ISOLDE, CERN, the most neutron-rich members of the Ba isotopic chain, ranging from $A=148$ to 150 , have been populated by β decay of their progenitors $^{148-150}\text{Cs}$. This is one of the regions of the nuclear chart where octupole deformations have been predicted to occur. The IDS setup includes fast $\text{LaBr}_3:\text{Ce}$ scintillator detectors to measure the level's lifetimes and extract information on the reduced electric quadrupole transition probability $B(E2:2^+ \rightarrow 0^+)$, which is strictly connected to the deformation of the nucleus. In the upper panel of Fig. 12.5, the newly measured transitions in ^{148}Ba and ^{150}Ba are reported in red, resulting in nearly degenerate spectra with interleaving levels characterized by positive and negative parity [38]. This is a signature of strong octupole correlations. The experimental findings are well described in a beyond-mean field approximation including symmetry restoration and configuration mixing within the generator coordinate method [39, 40]. This is shown in the bottom panel of Fig. 12.5, where the calculated Potential Energy Surfaces and collective wave functions are given.

12.1.5 Nuclear Astrophysics: The LUNA Experiment in the Gran Sasso Underground Laboratory

Nuclear astrophysics is based on the precise knowledge of cross sections of thermonuclear fusion reactions occurring during the Big Bang Nucleosynthesis and in all burning stages of different stellar objects. Such primary ingredients are also very important for stellar models and for describing the evolution of the Universe.

Thermonuclear fusion reactions occur in the energy range of the Gamow peak [41], arising from the convolution of the energy distribution of nuclei in the stellar plasma and the tunnelling probability through the Coulomb barrier between the interacting charged particles. For a central temperature of 1.5×10^7 K, as in our Sun, the Gamow peak for the p-p chain or the CNO cycle of the Hydrogen burning reaction is below 30 keV, leading to extremely low reaction cross sections, down to the femto-barn level. It follows that a direct investigation of thermonuclear reactions near their Gamow energy is often beyond technical capabilities, as the signal-to-noise ratio is severely dominated by any source of unwanted background.

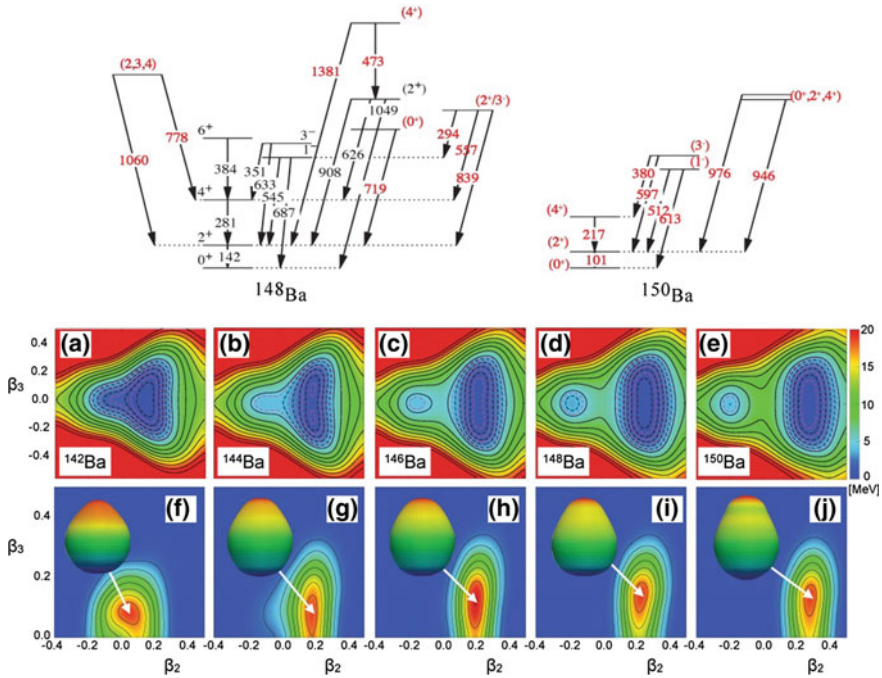


Fig. 12.5 Top panel: level schemes for ^{148}Ba (left) and ^{150}Ba (right) extracted from β - γ - γ coincidence analysis (newly found transitions in red). Bottom: Beyond mean field calculations predicting the evolution of the non axially deformed shapes in the Ba isotopic chain [38]

A drastic background reduction can be achieved by carrying out the measurements in a deep underground laboratory, such as under the Gran Sasso Mountain in Italy (LNGS). The rock overburden of about 1400 m (3800 m water equivalent) reduces the muon, neutron and gamma component of the cosmic background by a factor of 10^6 , 10^3 and 10, respectively, as compared to a laboratory on the Earth’s surface.

At LNGS, the LUNA collaboration performs measurements of astrophysical interest using a 400 kV accelerator [42]. Most important results are reviewed in [43–45]. As an example, Fig. 12.6 reports the first direct detection of the 70 keV resonance of the $^{17}\text{O}(p, \alpha)^{14}\text{N}$ reaction, very important for the nucleosynthesis in AGB stars [46].

The big step forward in the field will be given by the LUNA MV project, consisting in the installation of a new, 3.5 MV, accelerator inside the LNGS laboratory. The machine will be able to provide intense beams of H^+ , $^4\text{He}^+$, $^{12}\text{C}^+$ e $^{12}\text{C}^{++}$ in the energy range: 350 keV–3.5 MeV. The scientific program of LUNA MV foresees the measurement of a number of key reactions, such as (i) the $^{14}\text{N}(p, \gamma)^{15}\text{O}$ reaction (fundamental for the understanding of the solar composition), (ii) the $^{12}\text{C} + ^{12}\text{C}$ reaction (influencing the global chemical evolution of the Universe), and (iii) the so-called neutron source reactions, $^{13}\text{C}(\alpha, n)^{16}\text{O}$ and $^{22}\text{Ne}(\alpha, n)^{25}\text{Mg}$ (entering

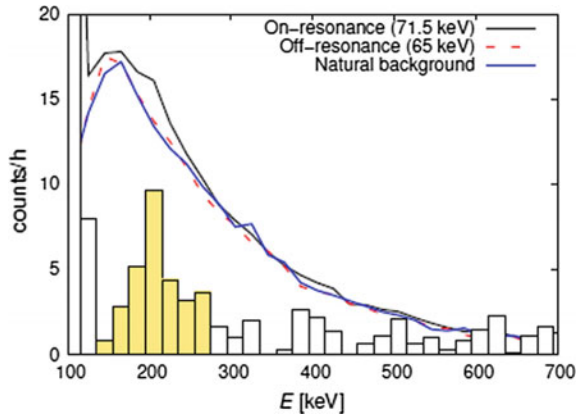


Fig. 12.6 On-resonance (proton energy of 71.5 keV), off-resonance and background spectra time normalized for the study of the 64.5 keV resonance of $^{17}\text{O}(p, \alpha)^{14}\text{N}$ [46]. The histogram (in arbitrary units) results from the difference between the on-resonance and the background spectrum

the nucleosynthesis of almost half of the heavy elements ($A > 60$) through the slow neutron capture process (s-process)).

12.2 Technical Developments

A significant part of the technical activities carried out in Milano within the Nuclear Structure group concerns research and developments for scintillators detectors, based on novel types of materials. The aim is the construction of high performance scintillator arrays for the research studies outlined in the previous sections. An example is the array composed of large size $3.5'' \times 8''$ crystals (now operational in Japan and previously employed coupled to the AGATA tracking array [2]), or the one consisting of $3'' \times 3''$ $\text{LaBr}_3:\text{Ce}$ (recently operational together with the HPGe GALILEO arrays at Legnaro National Laboratory).

In the last 10 years a large number of new high light-yield scintillator materials have, in fact, been discovered. The firsts, most famous, are the Lanthanum Halides, which have been the target of an intense R&D activity, providing the starting point for the design and development of several new highly performing $\text{LaBr}_3:\text{Ce}/\text{LaCl}_3:\text{Ce}$ based detectors arrays [47–50], such as HECTOR⁺ [47], PARIS [48] and CALIFA [49].

Apart from Lanthanum Halides, a large number of different new promising scintillators are now commercially available in small volumes, as for example CeBr_3 , CLYC, GYGAG:Ce, CLLB, CLLBC [51–53]. In addition, it was discovered that co-doping inorganic scintillators might have the effect of increasing the crystal proportionality, thus significantly improving the energy resolution [54, 55]. For example,

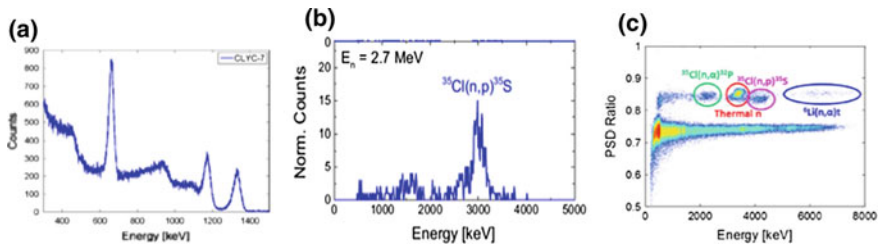


Fig. 12.7 Panel **a** the response of a $1'' \times 1''$ CLYC scintillator (enriched with ${}^7\text{Li}$) to ${}^{137}\text{Cs}$ and ${}^{60}\text{Co}$ sources. Panel **b** energy spectrum measured with a $1'' \times 1''$ CLYC scintillator when illuminated with monochromatic 2.7 MeV neutrons. Panel **c** Pulse Shape Discrimination (PSD) matrix of the $1'' \times 1''$ CLYC scintillator. Fast, thermal neutrons and γ rays correspond to a PSD level of 0.85 and 0.72, respectively [52]

it has been reported that a small sample of $\text{LaBr}_3:\text{Ce}$ co-Doped with Sr^+ provides an energy resolution of 2.3% at 662 keV, instead of the typical 3%. All these scintillator materials have an energy resolution which is much better than that of NaI, a good time resolution and some of them possess also Pulse Shape Discrimination (PSD) properties to identify the incident radiation (γ , protons, neutrons, alpha).

Since 2013, CLYC scintillators are available in very small samples, with an energy resolution slightly lower than 5% at 662 keV and the unique feature of identifying fast and thermal neutrons: a 1 cm thick sample of CLYC (enriched with ${}^6\text{Li}$) provides approximately 80% efficiency for thermal neutron detection and excellent Pulse Shape Discrimination between neutrons and γ -rays (similar or even better than ${}^3\text{He}$ detectors). In addition, thanks to ${}^{35}\text{Cl}$ nuclei, CLYC scintillators can also measure the kinetic energy of fast neutrons. Owing to a time resolution below 1 ns, CLYC detectors have the unique capability of identifying neutrons and measure their kinetic energy, using both the pulse height and Time Of Flight technique. Figure 12.7 shows the response of a CLYC enriched with ${}^7\text{Li}$ to gamma rays and monochromatic 2.7 MeV fast neutrons.

The CLLBC scintillator, composed by a molecule which is the union of the LaBr_3 and CLYC molecule, is available since few months only and it is expected to have an energy resolution comparable to that of $\text{LaBr}_3:\text{Ce}$ and Pulse Shape Discrimination properties very similar to those of CLYC. Tests are ongoing at the Milano Physics Department.

Finally, it is important to remark that all high light yield crystals can potentially provide good position sensitivity within continuous crystals (as it was shown with $\text{LaBr}_3:\text{Ce}$) [56, 57] (see Fig. 12.8). This feature could be used to reduce the Doppler Broadening effect in basic research and be extremely useful in several fields of applications, as for example in homeland security, medical-imaging or radiotherapy dosimetry.

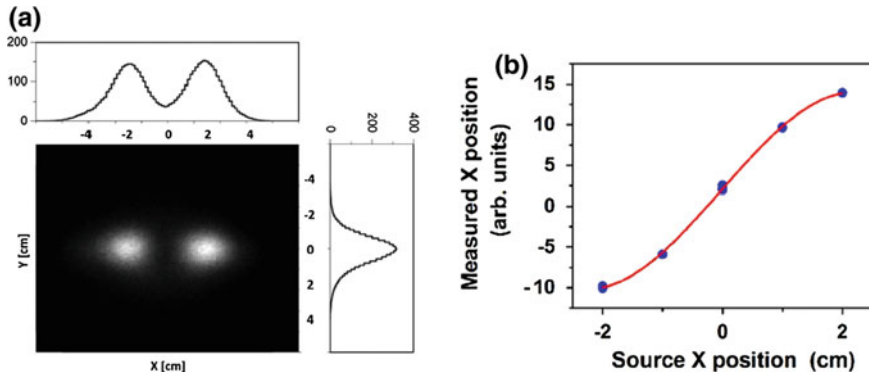


Fig. 12.8 Panel **a** two-dimensional image corresponding to two sources placed 4 cm apart in the x direction. On the top and on the right, the profiles along the x - and y -axes are shown, respectively. Panel **b** the spatial linearity measured in a $3'' \times 3''$ LaBr₃:Ce detector [56]

References

1. P. Moller et al., Phys. Rev. C **67**, 055802 (2003)
2. S. Akkoyun et al., Nucl. Instrum. Methods Phys. Res. A **668**, 26 (2012)
3. J.L. Wood et al., Phys. Rep. **215**, 101 (1992)
4. K. Heyde, J.L. Wood, Rev. Mod. Phys. **83**, 1467 (2011)
5. B. Singh et al., Nucl. Data Sheets **97**, 241 (2002)
6. A. Lopez-Martens et al., Prog. Part. Nucl. Phys. **89**, 137 (2016)
7. S. Leoni, A. Lopez-Martens, Phys. Scr. **91**, 063009 (2016)
8. P. Walker, J. Dracoulis, Nature **399**, 35 (1999)
9. S.M. Polikanov, Sov. Phys. Usp. **15**, 486 (1973)
10. J. Kantele et al., Phys. Rev. Lett. **51**, 91 (1983)
11. P. Butler et al., J. Phys. G **6**, 1165 (1980)
12. P. Bonche et al., Nucl. Phys. A **500**, 308 (1989)
13. M. Girod et al., Phys. Rev. Lett. **62**, 2452 (1989)
14. P. Möller et al., Phys. Rev. Lett. **103**, 212501 (2009)
15. B.P. Crider et al., Phys. Lett. B **763**, 108 (2016)
16. A.I. Morales et al., Phys. Rev. C **93**, 034328 (2016)
17. A.I. Morales et al., Phys. Lett. B **765**, 328 (2017)
18. Y. Tsunoda et al., Phys. Rev. C **89**, 031301 (2014)
19. S. Leoni, B. Fornal, N. Marginean et al., Phys. Rev. Lett. **118**, 162502 (2017)
20. M.B. Tsang et al., Phys. Rev. Lett. **102**, 062501 (2009)
21. P.F. Bortignon, A. Bracco, R.A. Broglia, *Giant Resonances: Nuclear Structure at Finite Temperature* (Harwood Academic, New York, 1998)
22. Y.F. Niu et al., Phys. Rev. Lett. **114**, 142501 (2015)
23. D. Montanari et al., Phys. Lett. B **697**, 288 (2011)
24. D. Montanari et al., Phys. Rev. C **85**, 044301 (2012)
25. D. Montanari et al., Eur. Phys. J. A **47**, 4 (2011)
26. G. Bocchi et al., Phys. Lett. B **760**, 273 (2016)
27. G. Colò, P.F. Bortignon, G. Bocchi, Phys. Rev. C **95**, 034303 (2017)
28. R.A. Broglia, G. Colò, G. Onida, H.E. Roman, *Solid State Physics of Finite Systems* (Springer, Germany, 2004)
29. M.N. Harakeh, A. van der Woude, *Giant Resonances: Fundamental High-Frequency Modes of Nuclear Excitation*, Oxford Science Publications (2001)

30. O. Wieland, A. Bracco, F. Camera et al., Phys. Rev. Lett. **102**, 092502 (2009)
31. A. Bracco, F.C.L. Crespi, E.G. Lanza, Eur. Phys. J. A **51**, 99 (2015) and ref. therein
32. F.C.L. Crespi et al., Phys. Rev. Lett. **113**, 012501 (2014)
33. L. Pellegrini et al., Phys. Lett. B **738**, 519 (2014)
34. M. Krzysiek et al., Phys. Scr. **89**, 054016 (2014)
35. F.C.L. Crespi et al., Phys. Rev. C **91**, 024323 (2015)
36. L. Pellegrini, A. Bracco et al., Phys. Rev. C **92**, 014330 (2015)
37. *Extreme Light Infrastructure—Nuclear Physics (ELI-NP)*. <http://www.eli-np.ro>
38. R. Lica et al., J. Phys. G: Nucl. Phys. **44**, 054002 (2017); Phys. Rev. C **97**, 024305 (2018)
39. M. Yao, E. Zhou, Z. Li, Phys. Rev. C **92**, 041304 (2015)
40. R.N. Bernard, L.M. Robledo, T.R. Rodriguez, Phys. Rev. C **93**, 061302(R) (2016)
41. C. Iliadis, *Nuclear Physics of Stars*, Wiley Weinheim (2007)
42. A. Formicola et al., Nucl. Instr. Meth. A **507**, 609 (2003)
43. H. Costantini et al., Rep. Prog. Phys. **72**, 086301/1 (2009)
44. C. Broggini et al., LUNA: Ann. Rev. Nucl. Part. Sci. **60**, 53 (2010)
45. Topical Issue, Underground nuclear astrophysics and solar neutrinos: Impact on astrophysics, solar and neutrino physics. Eur. Phys. J A **52** (2016)
46. C. Bruno et al., Phys. Rev. Lett. **117**, 142502 (2017)
47. F. Camera et al., EPJ Web Conf. **66**, 11008 (2014)
48. A. Maj et al., Acta Phys. Pol. B **40**, 565 (2009)
49. E. Nacher et al., Nucl. Instr. Meth. A **769**, 105 (2015)
50. A. Giaz et al., Nucl. Instr. Meth. A **729**, 910 (2013)
51. A. Giaz et al., Nucl. Instr. Meth. A **804**, 212 (2015)
52. A. Giaz et al., Nucl. Instr. Meth. A **810**, 132 (2016); Nucl. Instr. Meth. A **825**, 51 (2016)
53. R. Hawrami et al., IEEE Trans. Nucl. Sci. **63**, 509 (2016)
54. K. Yang et al., IEEE Trans. Nucl. Sci. **63**, 416 (2016)
55. M.S. Alekhin et al., Appl. Phys. Lett. **102**, 161915 (2013)
56. N. Blasi et al., Nucl. Instr. Meth. A **839**, 23 (2016)
57. A. Giaz et al., Nucl. Instr. Meth. A **772**, 103 (2015)

Chapter 13

Planet Formation in the ALMA Era



**Giuseppe Lodato, Benedetta Veronesi, Maria Giulia Ubeira Gabellini,
Enrico Ragusa, Alessia Franchini, Giovanni Dipierro,
Hossam Aly and Daniel J. Price**

Abstract The field of planet formation has been transformed recently by the completion of high-resolution imaging techniques in the infrared (SPHERE) and mm-wavelength bands (ALMA) that allow us to obtain images at unprecedented high angular resolution of protostellar discs, that are the site of planet formation. At the same time, large planet-hunting campaigns have been carried out in the last decades, showing us the complexity and the diversity of exo-planetary systems. The architecture of planetary systems is shaped by a variety of mechanisms: from the hydrodynamical interactions between the various components of a protostellar disc (gas, dust, proto-planets), to the N -body interactions between newly formed planets. In this contribution, I will describe the efforts carried out at the University of Milan to understand the hydrodynamics of protostellar discs and how such dynamical mechanisms determine the structures observed with high-resolution imaging. This is done mostly by performing high-resolution numerical simulations of the coupled evolution of gas and dust around a newly born star.

G. Lodato (✉) · B. Veronesi · M. G. Ubeira Gabellini · E. Ragusa · A. Franchini · H. Aly
Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Via Celoria 16,
Milan, Italy
e-mail: giuseppe.lodato@unimi.it

M. G. Ubeira Gabellini
European Southern Observatory (ESO), Garching, Germany

G. Dipierro
Department of Physics and Astronomy, University of Leicester,
University Road, Leicester, UK

D. J. Price
Monash Centre for Astrophysics (MoCA) and School of Physics and Astronomy,
Monash University, Melbourne, Australia

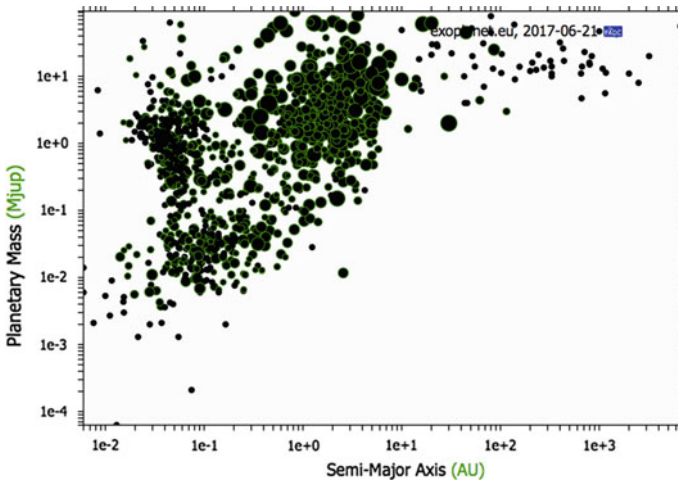


Fig. 13.1 Distribution of planetary masses and semi-major axis of the currently confirmed ~ 1000 exo-planets. The size of the points is proportional to the orbital eccentricity. One can notice the presence of the so-called “hot Jupiter” population at a separation of ~ 0.1 au and masses of the order of Jupiter, and the population of directly imaged planets at tens to hundreds of au from their host star. Data from the Extrasolar Planets encyclopedia: <http://exoplanet.eu>

13.1 Introduction

Planet formation is one of the most exciting topics in contemporary astrophysics. The excitement mostly comes from two different observational and instrumentation developments. On the one hand, large surveys, both from the ground (through Doppler measurement of photospheric stellar lines) and from space (through missions dedicated to the detection of stellar transits, such as Kepler¹ and the upcoming European mission PLATO²), have as main objective the discovery of extra-solar planets, that are planets orbiting stars different than the Sun. On the other hand, extremely high resolution has been achieved both at infrared wavelengths, with the SPHERE instrument at the Very large Telescope (VLT), and at sub-mm wavelengths with the Atacama Large Millimeter Array (ALMA) interferometer. In both cases, these instruments are able to probe nearby protostellar discs at the spatial resolution of a few astronomical units (au), corresponding to an angular resolution of a few milliarcseconds (mas), thus probing directly the regions of planet formation.

For what concerns the currently known exo-planet population, we have currently discovered over a thousand planets (with over three thousand candidates). We have certainly gone a long way from the discovery of the first planet orbiting a solar-type star in 1995, with the discovery of 51 Peg b [1]. The planet population shows a few remarkable properties, see Fig. 13.1. Firstly, there are a large number of planets, with

¹https://www.nasa.gov/mission_pages/kepler/main/index.html.

²<http://sci.esa.int/plato/>.

a mass close to Jupiter, orbiting their stars at distances much closer than the distance of Mercury from the Sun. This population of “hot Jupiters”, as they are called, is unlikely to have formed in situ, since there is not enough gas at those distances to the star in a typical protostellar disc. Their presence thus is a clear indication that planet migration must occur during the first stages of planet formation and evolution. Such planet migration can occur due to torques exerted between the young planet and the surrounding gaseous discs [2] (expected to result in a small planet eccentricity) or through planet-planet scattering after the gaseous disc has been accreted by the central star [3] (expected to produce significant orbital eccentricities). Secondly, one can easily see in the Figure the typical triangular distribution of planets in the mass-semimajor axis plots produced by observational biases. At distances of a few au, the main discovery technique is the so-called “radial velocity” technique, whereby a planet is discovered by the reflex motion that it induces on the star, resulting in a periodic Doppler shift of the stellar photospheric lines. The limited spectroscopic resolution of these experiments naturally produces a bias in favour of more massive planets orbiting closer to the star, so that at a larger distance only the most massive planets can be detected (with the minimum detectable planet mass scaling as $M_{\min} \propto \sqrt{a}$, where a is the semi-major axis). Thirdly, at very large distances from the star (tens to hundred au) one can see the small but interesting population of directly imaged planets, such as HR8799 [4]. The currently leading planet formation theory—the so-called core accretion model [5]—predicates a gradual growth of planets from microscopic dust into large planetesimals, that eventually form an Earth mass planetary embryo. Once the embryo reaches a mass of several Earth masses, it is able to accrete a large gaseous envelope and become a Jupiter mass planet. Clearly this model, being composed of several steps, is quite slow, and to assemble a giant planet, typically a timescale of the order of a few Myrs is needed. This is dangerously close to the typical lifetime of protostellar discs, so that there is the risk that, once planets are finally massive enough to accrete the gaseous envelope, the gas has already been dispersed or accreted by the central star. The population of giant planets at tens to hundreds of au poses a strong challenge to planet formation theories, since the formation timescale becomes excessively long at these distances to be feasible.

The planet-hunting campaigns offer an insight of the outcome of planet formation, since they target fully formed main sequence stars. If we want to get an insight of planet formation as it happens, we need to target younger stars (the so-called T Tauri stars) whose emission is dominated at long wavelengths by the circumstellar discs that are the site of planet formation. Such discs are composed of gas and dust (with a gas/dust ratio of approximately 100, as observed in the Interstellar Medium), and have masses of the order of 10^{-4} – $10^{-2} M_{\odot}$ [6]. The recent completion of ALMA and of SPHERE has allowed us for the first time to obtain accurate imaging of such discs at au lengthscales, where planet formation takes place. Such images have revealed an incredibly rich diversity of morphologies, showing ringed structures (as in the case of HL Tau [7]), spirals (as in HD135344B [8], and in Elias 2-27 [9]), and horseshoe structures (as in the case of HD142527 [10]). All such sub-structures are the likely result of dynamical interactions between the gas and dust in the disc and newly forming planets.

It is important to stress that gas and dust in the disc are coupled by drag. This is typically in the form of a Stokes or an Epstein drag. The coupling can be characterized in terms of a stopping time, t_s , such that

$$\frac{d\mathbf{v}_{\text{dust}}}{dt} = -\frac{\mathbf{v}_{\text{dust}}}{t_s}, \quad (13.1)$$

where \mathbf{v}_{dust} is the dust velocity relative to the gas. The effectiveness of coupling depends on the relative magnitude of t_s and the orbital timescale Ω^{-1} (where Ω is the angular velocity of circular orbits around the central star), so that one generally uses the dimensionless Stokes number to characterize the dust dynamics, defined as:

$$St = \Omega t_s = \frac{\rho_d s}{\Sigma_g}, \quad (13.2)$$

where ρ_d is the internal dust density, s is the effective dust grain size and Σ_g is the surface density of the gaseous disc and where the last equality holds for the most typical case of Epstein drag. One thus sees that gas-dust coupling is proportional to the dust size, with larger dust being less coupled to the gas. Dust with $St \ll 1$ is strongly coupled, while dust with $St \gg 1$ is loosely coupled.

13.2 Hydrodynamical Simulations of Protostellar Discs

At the Physics Department at the University of Milano the approach to protostellar disc modeling is mostly related to high resolution numerical simulations of the system comprising one (or more) stars, a gaseous and a dusty disc and one (or more) planets embedded in the disc. As an outcome, the simulations provide us with the 3D density distribution of the gas and of the dust (with different grain sizes), as well as the gas temperature. We then use such distributions as an input to Monte Carlo radiative transfer code to provide us with the dust temperature and emissivity at the required wavelengths. These “infinite resolution” images are then processed to simulate the instrumental response and compared to the actual observations. In this way we can tune the parameters of our simulations to match the observations in terms of a physically motivated model.

We solve the equations of hydrodynamics in 3D using the code PHANTOM [11]. PHANTOM³ is a Smoothed Particle Hydrodynamics (SPH) code, developed by Daniel Price in Monash (Australia) and of which the Department of Physics in Milano is a developing node. It is freely available under the GPLv3 licence. SPH is a Lagrangian algorithm where a fluid is discretized into mass elements (rather than volume elements, as done in Eulerian codes), called “particles”, whose properties are obtained through a weighted average over neighbouring particles using a weighting kernel

³<https://phantom.sph.bitbucket.io>.

that falls off with distance on a typical scale, called the “smoothing length”, that is proportional to the mean interparticle separation. In this way, high density regions are naturally evolved with a very high spatial resolution.

PHANTOM has been introduced from the very beginning with applications to accretion disc dynamics [12] and turbulence in 3D [13]. It is a highly efficient, modular code, parallelized with OpenMP with excellent scaling properties and MPI. It includes the most advanced algorithm for SPH: full Lagrangian derivation, with adaptive smoothing and adaptive softening, with the most advanced artificial viscosity switches [14], as well as artificial thermal conduction. Implemented physics includes the fluid self-gravity, computed using a tree. Dust is modelled either with separate sets of dust and gas particles [15], appropriate for $St \gg 1$, where the dust and gas are treated as two interacting fluids, or as a single fluid mixture modelled on a single set of SPH particles [16], appropriate for small St , where the SPH particles represent the whole gas-dust mixture and the dust fraction is evolved as a local property of the mixture.

Apart from its application to protostellar discs, PHANTOM is a versatile code that can be used on a variety of astrophysical environments. It includes a magneto-hydrodynamics (MHD) module. For applications to black hole systems, it includes a number of general relativistic effects, treated approximately within a pseudo-Newtonian approach. In particular, it includes Einstein’s apsidal precession (using a modified Newtonian potential [17]), Lense-Thirring precession around a spinning black hole [18], and a prescription for the decay of a black hole binary system due to the emission of gravitational waves [19].

In order to simulate a protostellar disc over a reasonable dynamical range with enough resolution, the typical number of particles to be used is of the order of a few millions. On a 16-core machine, a typical simulation lasts for a few weeks of wall-time.

13.3 A Few Applications

In this section we describe a number of applications of our modeling to explain some of the most striking features observed in high resolution images of protostellar discs.

13.3.1 *Modeling HL Tau*

The HL Tau disc was the first disc to be observed with the ALMA most extended baseline. Its image was greeted with enthusiasm and shock by the community. Indeed, its spectacular series of dark and bright rings, see Fig. 13.2, reminiscent of Saturn rings, pointed to the presence of newborn Saturn mass planets. HL Tau is a young T Tau star, with an age of roughly one million years and a mass of $\sim 1.3M_{\odot}$, located in the Taurus-Auriga star forming region at 140 pc from the Sun.

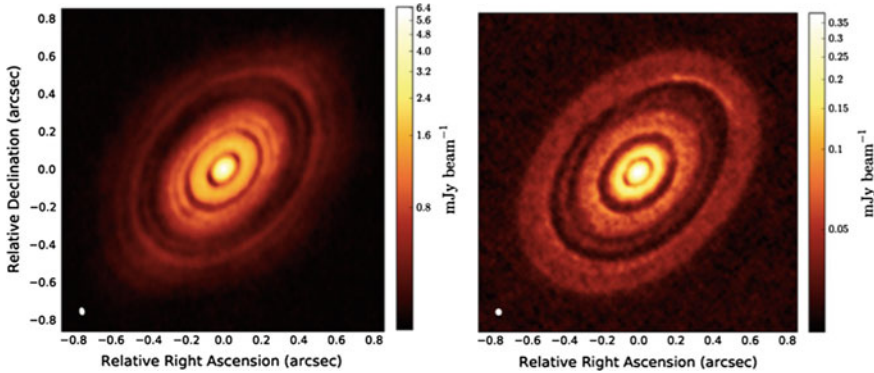


Fig. 13.2 Left: image of the HL Tau disc obtained with the ALMA interferometer at 1.3 mm [7]. Right: Synthetic ALMA image of one of our simulated discs [20]

Initially, theorists were puzzled by the axisymmetric appearance of the disc. Indeed, while it is well known that the presence of massive planets in a disc produces deep gaps in the surface density, it is also expected that a prominent spiral structure would form. This puzzle was solved by our group [20], by showing that the gas and the dust respond differently to the presence of a planet. While the gas, that supports the development of density waves, that propagate through the action of pressure, responds to the planet presence by launching spiral density waves, the dust, that is pressureless, tends to produce axisymmetric gaps. Obviously, this phenomenon is dependent on the size of dust grains, so that small grains (more coupled to the gas) follows the gas motion and produces spirals, while large grains display more of a ring like morphology. We performed our simulations by using 6 different grain size, from micron to ten cm sizes. By interpolating the emission maps and assuming a given dust size distribution, we produced a synthetic disc emission map that we then processed through the instrumental response to finally obtain a synthetic image. Our best match to the observations in terms of gap sizes, location and contrast is obtained with three planets with masses of $0.2M_{\text{Jupiter}}$, $0.27M_{\text{Jupiter}}$ and $0.55M_{\text{Jupiter}}$, orbiting the central star at distances equal to 13.2 au, 32.3 au and 68.8 au, respectively.

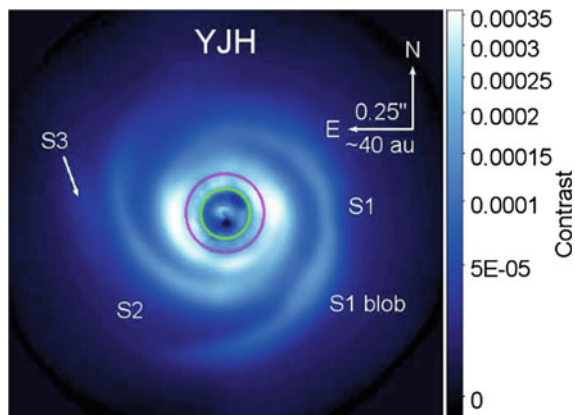
Our model is currently regarded as a standard model to explain the HL Tau system. It also poses several interesting questions: the prevalent model for planet formation—the core accretion model—as mentioned above is a relatively slow process, taking roughly several million years to form a Jupiter-like planet. In our model, on the other hand, HL Tau is surrounded by three Jupiter mass planets at its very young age, thus posing a challenge for the core accretion model. It is also worth mentioning that a strikingly similar structure is also observed in another protostellar disc, the one around TW Hya [21], that has a very different age of 10 Myr. How two objects of such different age, one at the beginning and the other at the end of the disc lifetime, can share so similar morphology is still an open question.

13.3.2 Modeling HD135344B and CQ Tau

Another spectacular morphology observed in protostellar disc is the large scale spiral structure in HD135344B, see Fig. 13.3 [22]. HD135344B is a pre-main sequence star with a mass of $1.7M_{\odot}$, at a distance of 157 pc, surrounded by an almost face-on disc. The disc has a relatively low mass $\sim 0.01M_{\odot}$ (although estimates of the disc mass is subject to not fully understood systematic uncertainties [23]), which then appears to exclude a self-induced origin of the spiral structure as a consequence of a gravitational instability in the disc (akin to the case of spiral galaxies [24]).

The emission in the infrared is dominated by scattering of stellar light off the disc surface by micron-sized dust. At these small sizes, the dust is expected to be strongly coupled to the gas and so one possible explanation of the observed spiral structure can lie in the same mechanism that was originally proposed for HL Tau, that is the development of density waves due to the presence of an unseen planetary mass companion. It was initially suggested that the two arms of the observed spiral were produced by two low mass planets (with masses significantly below Jupiter), each inducing one arm [25]. Another possible scenario is one were a single massive planet is responsible for both arms [26]. We have applied the analysis outlined above for HL Tau also to the case of HD135344B (Veronesi et al., in prep.) and we have been able to exclude the two-planet scenario, that would produce a completely different contrast between the arm and inter-arm regions with respect to the observations. The single-planet scenario is more consistent with the observed spiral but however a further planet is still needed at small separations to explain the observed inner ring in the scattered-light image. Our best fitting model to HD135344B currently uses an outer planet with a mass of $6M_{\text{Jupiter}}$ located at 140 au from the star and an inner, $4M_{\text{Jupiter}}$ planet, located at 35 au. This model reproduces fairly well both the infrared and radio observations, but a further exploration of the available parameter space is needed.

Fig. 13.3 Infrared image of the HD135344B disc and its large scale spiral structure obtained with SPHERE [22]



So far, we have mostly discussed dust diagnostics for protostellar discs either in large, mm-sized grains (through ALMA) or in micron-sized grains that emit in the infrared (SPHERE). The small grains, being strongly coupled to the gas, are usually assumed to be a good, although indirect, gas tracer. ALMA is also able to provide direct gas tracers through the observations of specific line transitions. Typical molecular lines that fall in the ALMA bands are the lines of carbon monoxide in its various isotopologues: ^{12}CO , ^{13}CO and C^{18}O . Generally, modeling the emission in such gas species presents the additional difficulty in having to solve for the chemical structure of the disc, in addition to its physical properties. A complex chemical network is already computationally expensive and it is currently unfeasible to couple it with another computationally challenging tool such as an SPH simulation. Our group and its collaborators have recently obtained CO and dust data through ALMA for another interesting object, CQ Tau, a transition disc (see below) for which our infrared images (obtained with SPHERE and the Large Binocular Telescope Interferometer) suggest the presence of a *directly imaged* planet (in contrast to the previous cases, where the planetary presence was only inferred based on its effects on the disc morphology). We are currently in the process of modeling the ALMA data in terms of simplified, axisymmetric disc models using a radiative transfer and chemical network tool called DALI [27], see Fig. 13.4 (Ubeira Gabellini et al., in prep.). Once this modeling is

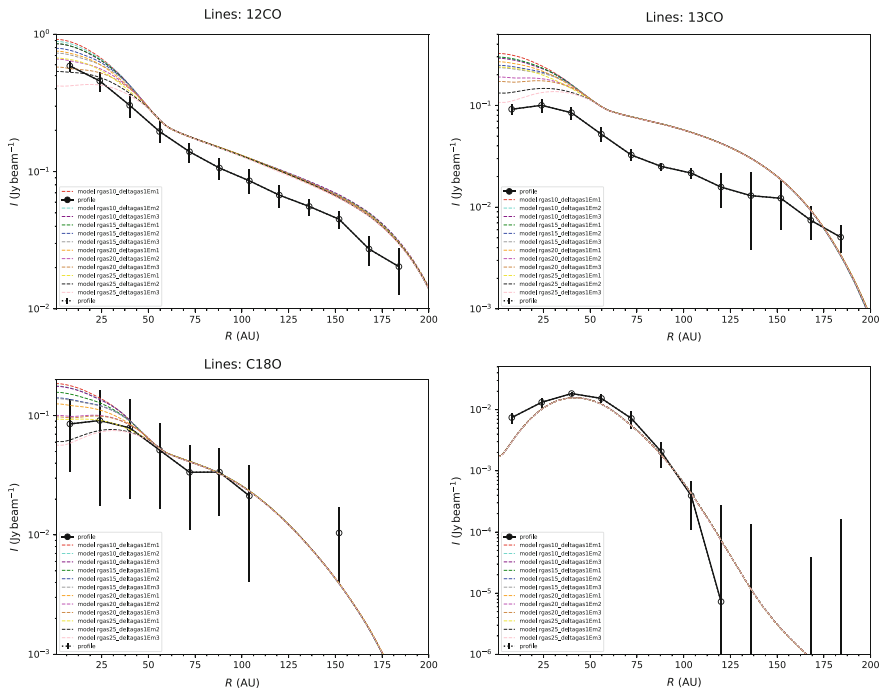


Fig. 13.4 Our best fit model to the radial emissivity profile in CQ Tau in different diagnostics: ^{12}CO (upper left), ^{13}CO (upper right), C^{18}O (lower left) and dust continuum (lower right)

complete, we will proceed to perform full 3D hydrodynamics simulations of the system including a planet, in order to put further constraints on the purported directly imaged planet.

13.3.3 Horseshoes in Transition Discs

Transition discs are a class of protostellar discs characterized by strong emission at mm wavelengths and little or no emission in the infrared. In the pre-ALMA era, when discs were unresolved and only integrated Spectral Energy Distributions (SED) were available, this had been interpreted in terms of the presence of a large cavity devoid of emitting material. Such interpretation has been spectacularly confirmed by ALMA imaging. Moreover, it also has been understood that such large dust cavities (dust being the main emitter in the disc) do not necessarily correspond to gas cavities, that can still fill up the inner regions of the disc. The origin of the cavity is traditionally attributed again to the presence of a massive planet or a stellar companion. The different cavity size in gas and dust is yet another confirmation of the different response of the two components to dynamical perturbations. One of the most striking features of some transition discs imaged with ALMA is the fact that the dust structure at the cavity edge can be very asymmetric, displaying a prominent horseshoe morphology. This is the case of the prototypical IRS48 disc [28], or of the large crescent seen in HD142527 [10]. These horseshoes have often been interpreted in terms of vortices generated at the edge of the cavity by the Rossby-wave instability [29]. However, in order to be long lived, such vortical structures require that the disc is almost inviscid, contrary to the evidence that accretion—usually associated with fairly high level of viscosity—is present in most transition discs.

We [30] have recently proposed a novel explanation for the formation of horseshoes in transition discs, whereby they are the manifestation of a dynamical instability associated with gas streams within the cavity in a circumbinary disc. If the mass ratio of the two components of the binary is large enough (above ~ 0.04) the disc becomes eccentric and naturally produces a stable horseshoe overdensity, which is long lived also in the presence of high viscosity, see Fig. 13.5. It is worth pointing out that one of the best examples of such horseshoes, HD142527, is indeed known to be a binary system, with a mass ratio consistent with our expectations [31].

13.3.4 A Few Application to Black Hole Discs

To conclude this overview of the recent activities in this field at the University of Milano, we would like to mention some of the results that have been obtained in the context of accretion discs around black holes.

One of the most exciting recent discoveries in the field of black hole growth has been in the context of transient phenomena. In particular, large attention has been

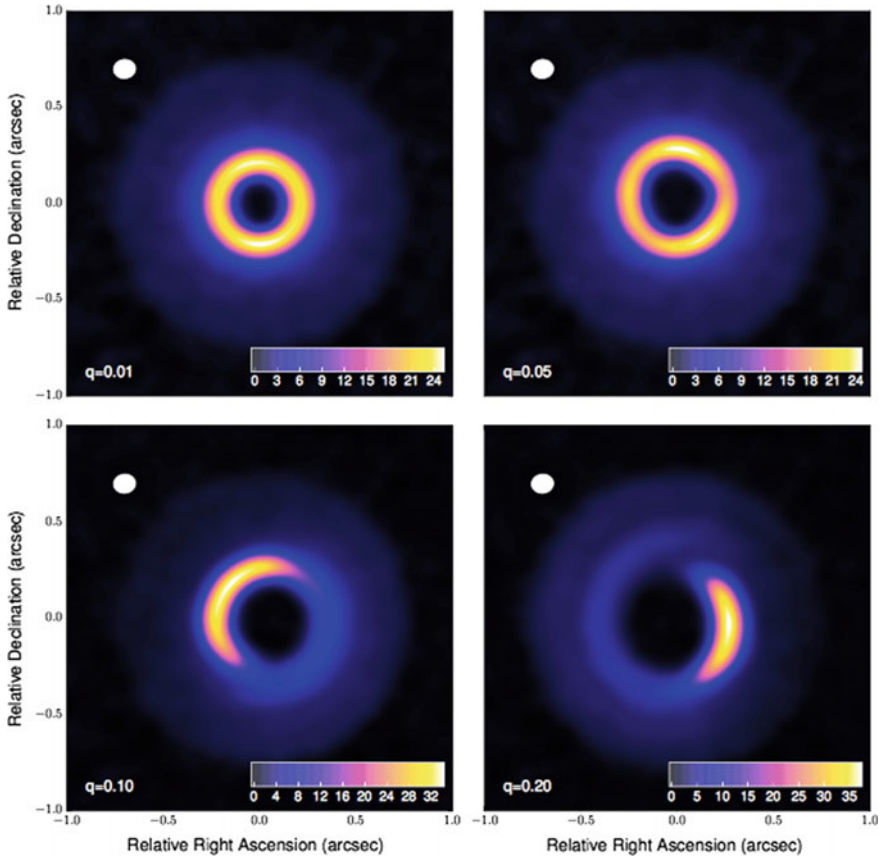


Fig. 13.5 ALMA synthetic observations at 345 GHz of some of our simulated circumbinary discs, with mass ratios of $q = 0.01$ (upper left panel), $q = 0.05$ (upper right), $q = 0.1$ (lower left) and $q = 0.2$ (lower right). It can be clearly seen that for the highest mass ratio a prominent horseshoe structure naturally develops and is detectable with ALMA. From [30]

devoted to the study of Tidal Disruption Event, where a star gets so close to a black hole to be torn apart from the black hole's tidal field. The distance at which disruption occurs, called the tidal radius, is given by $R_t = (M_*/M_{\text{BH}})^{1/3} R_*$, where M_* and R_* are the stellar mass and radius, respectively, and M_{BH} is the black hole mass. If a solar type star is disrupted by a $10^6 M_\odot$ black hole, such as the supermassive black holes (SMBH) residing in galaxy centers, the tidal radius lies at ~ 20 Schwarzschild radii from the hole. Once the star is disrupted, its debris fall back towards the hole and are rapidly accreted, giving rise to a luminous flare (with luminosities above 10^{44} ergs/s) lasting for months to years and with a typical lightcurve declining with time as $t^{-5/3}$ [32]. Such flares can be used as a tool to discover otherwise quiescent SMBH residing in the nucleus of normal galaxies. The expected rate of these events is one per 10,000 years per galaxy, and large survey have recently discovered dozens such

events. Due to the proximity of the event to the black hole horizon, it is expected that general relativistic effects will be important. We have recently proposed that the Lense-Thirring precession associated with the dragging of reference frames around a spinning black hole can cause the torus formed by the debris to precess rigidly around the SMBH, giving rise to quasi-periodic oscillations with a timescale of a few days [33]. Such oscillations, if detected, can offer a dynamical way to measure the spin of the black hole, an essential property that also determines the hole's efficiency in converting matter into luminosity. Such QPOs have also been observed around stellar mass black holes in X-ray binaries, for which we have also applied our rigid precession model to measure the black hole spin [34, 35].

Finally, a further line of research involves the evolution of warped accretion discs [12]. Such warped configurations can occur due to Lense-Thirring torques (see above) but also due to the presence of a central binary system, that can be either a stellar binary [36, 37] or a SMBH binary on its way to coalescence due to gravitational wave emission. If the binary is eccentric, a curious phenomenon can happen, whereby the circumbinary disc, rather than aligning to the orbital plane of the binary, finds a stable configuration by sitting in a polar orbit, perpendicular to the binary orbit [38]. Such configurations may be important both in the context of SMBH binaries but also in the context of protostellar binaries (such as the case of the GG Tau, where a misaligned disc orbits quite an eccentric binary [39, 40]).

13.4 Conclusions

The field of planet formation is being revolutionized in two different ways. A first revolution has already happened, and relates to the discovery of a large number of extra-solar planets, that finally allow us to compare the properties of our own solar system with a statistically significant (although not bias-free) sample of planetary systems around other stars. Such exoplanets detections provide us with an insight on the *outcome* of planet formation. A second revolution, that is happening now, is related to the process of planet formation *as it happens*, and has been made possible by the completion of high angular resolution facilities at long wavelengths, such as SPHERE in the infrared and especially ALMA, in the mm range. The new spectacular images of protostellar discs, the cradle in which planets are born, show a wealth of substructures that were unthinkable only a few years ago and require a deep re-thinking of our fundamental ideas on the evolution of such discs, that are a mixture of dust and gas, coupled by a drag force.

Here at the University of Milano, we work on the theoretical modeling of such systems, using some of the most advanced algorithms to treat the coupled hydrodynamics of gas and dust in the presence of gravitating point-masses (representing the young star and possibly newborn planets). We are among the developers of the code PHANTOM, a 3D Smoothed Particle Hydrodynamics code. In this paper, I have presented some of the applications of our modeling both with the aim to explain specific observed systems, such as HL Tau, HD135344B and CQ Tau (and thus to

put constraints on their physical parameters in terms of number, masses and location of young, yet unobserved, planets) and with the more general aim to discover and analyse new physical processes, such as those related to the formation of horseshoe asymmetries in circumbinary discs.

This is the time to put forward new models for the evolution of protostellar discs, models able to account for the complexities of the observed sub-structures and that will probably have to be based on a chaotic, turbulent scenario to determine the initial conditions and evolution of the star forming environment. Our work is rooted on such premises and aims to develop the next generation models of protostellar discs, beyond the traditional axisymmetric models first developed in the late '70s.

References

1. M. Mayor, D. Queloz, *Nature* **378**, 355 (1995)
2. D.N.C. Lin, J. Papaloizou, *MNRAS* **186**, 799 (1979)
3. F.A. Rasio, E.B. Ford, *Science* **274**, 954 (1996)
4. C. Marois, B. Macintosh, T. Barman, B. Zuckerman, I. Song, J. Patience, D. Lafrenière, R. Doyon, *Science* **322**, 1348 (2008)
5. R. Helled, P. Bodenheimer, M. Podolak, A. Boley, F. Meru, S. Nayakshin, J.J. Fortney, L. Mayer, Y. Alibert, A.P. Boss, *Protostars and Planets VI* (2014), pp. 643–665
6. G. Lodato, C.E. Scardoni, C.F. Manara, L. Testi, *MNRAS*, **472**, 4700 (2017)
7. A. Partnership, *ApJ Lett.* **808**, L3 (2015)
8. A. Garufi, S.P. Quanz, H. Avenhaus, E. Buenzli, C. Dominik, F. Meru, M.R. Meyer, P. Pinilla, H.M. Schmid, S. Wolf, *A&A* **560**, A105 (2013)
9. L.M. Pérez et al., *Science* **353**, 1519 (2016)
10. S. Casassus et al., *ApJ* **812**, 126 (2015)
11. D.J. Price, et al., *PASA*, **35**, e031 (2018)
12. G. Lodato, D.J. Price, *MNRAS* **405**, 1212 (2010)
13. D.J. Price, C. Federrath, *MNRAS* **406**, 1659 (2010)
14. L. Cullen, W. Dehnen, *MNRAS* **408**, 669 (2010)
15. G. Laibe, D.J. Price, *MNRAS* **420**, 2345 (2012)
16. G. Laibe, D.J. Price, *MNRAS* **440**, 2147 (2014)
17. E. Tejeda, S. Rosswog, *MNRAS* **433**, 1930 (2013)
18. R.P. Nelson, J.C.B. Papaloizou, *MNRAS* **315**, 570 (2000)
19. A. Cerioli, G. Lodato, D.J. Price, *MNRAS* **457**, 939 (2016)
20. G. Dipierro, D. Price, G. Laibe, K. Hirsh, A. Cerioli, G. Lodato, *MNRAS* **453**, L73 (2015)
21. S.M. Andrews, D.J. Wilner, Z. Zhu, T. Birnstiel, J.M. Carpenter, L.M. Pérez, X.N. Bai, K.I. Öberg, A.M. Hughes, A. Isella, L. Ricci, *ApJ Lett.* **820**, L40 (2016)
22. A.L. Maire et al., *A&A* **601**, A134 (2017)
23. A. Miotello, E.F. van Dishoeck, M. Kama, S. Bruderer, *A&A* **594**, A85 (2016)
24. G. Bertin, C.C. Lin, *Spiral Structure in Galaxies: A Density Wave Theory* (MIT Press, Cambridge, 1996)
25. T. Stolker et al., *A&A* **595**, A113 (2016)
26. R. Dong, J. Fung, *ApJ* **835**, 146 (2017)
27. S. Bruderer, N. van der Marel, E.F. van Dishoeck, T.A. van Kempen, *A&A* **562**, A26 (2014)
28. N. van der Marel, E.F. van Dishoeck, S. Bruderer, S.M. Andrews, K.M. Pontoppidan, G.J. Herczeg, T. van Kempen, A. Miotello, *A&A* **585**, A58 (2016)
29. S. Ataiee, P. Pinilla, A. Zsom, C.P. Dullemond, C. Dominik, J. Ghanbari, *A&A* **553**, L3 (2013)
30. E. Ragusa, G. Dipierro, G. Lodato, G. Laibe, D.J. Price, *MNRAS* **464**, 1449 (2017)

31. S. Lacour et al., *A&A* **590**, A90 (2016)
32. M.J. Rees, *Nature* **333**, 523 (1988). <https://doi.org/10.1038/333523a0>
33. A. Franchini, G. Lodato, S. Facchini, *MNRAS* **455**, 1946 (2016)
34. A. Franchini, S.E. Motta, G. Lodato, *MNRAS* **467**, 145 (2017)
35. S.E. Motta, A. Franchini, G. Lodato, G. Mastroserio, *MNRAS*, **473**, 431 (2018)
36. S. Facchini, G. Lodato, D.J. Price, *MNRAS* **433**, 2142 (2013)
37. G. Lodato, S. Facchini, *MNRAS* **433**, 2157 (2013)
38. H. Aly, W. Dehnen, C. Nixon, A. King, *MNRAS* **449**, 65 (2015)
39. P. Cazzoletti, L. Ricci, T. Birnstiel, G. Lodato, *A&A* **599**, A102 (2017)
40. H. Aly, G. Lodato, P. Cazzoletti, *MNRAS*, **480**, 4738 (2018)

Chapter 14

Generation and Active Control of Coherent Structures in Partially-Neutralized Magnetized Plasmas



**Giancarlo Maero, Roberto Pozzoli, Massimiliano Romé, Beatrice Achilli,
Nicola Cantini, Emilio Villa, Francesco Cavaliere and Daniele Viganò**

Abstract Penning-Malmberg (electro-magnetostatic) traps represent a great laboratory for the investigation of collective phenomena in plasmas and fluids, e.g., two-dimensional fluid dynamics and turbulence. As such they are usually exploited to trap single-species plasmas, which facilitates the diagnostics and manipulation of their dynamics as well as equilibrium states. Yet in some of the most advanced and challenging applications, such as sympathetic cooling, formation of ordered structures (crystals) or neutral antimatter, simultaneous confinement of multiple species, possibly with opposite sign of charge, is required, which complicates the plasma evolution and enhances instability mechanisms. We review a collection of experimental investigations focusing on the generation of an electron plasma by means of a radio-frequency (RF) electric field. The in-trap generation scheme also implies the presence of positive ions, and the complex dynamics of the two plasma components in the presence of a relatively strong RF drive gives rise to a wealth of new features with respect to single-species plasmas, most notably non-trivial equilibrium states coming out of the insurgence of long-lived coherent structures, or unexpected response to conventional excitation and manipulation schemes. These properties may be used to the aim of an active control of the electron sample positioning and charge.

G. Maero (✉) · R. Pozzoli · M. Romé · B. Achilli · N. Cantini · E. Villa
F. Cavaliere · D. Viganò
Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano,
Via Celoria 16, 20133 Milan, Italy
e-mail: giancarlo.maero@unimi.it

G. Maero · R. Pozzoli · M. Romé
INFN Sezione di Milano, Via Celoria 16, 20133 Milan, Italy

© Springer Nature Switzerland AG 2018
P. F. Bortignon et al. (eds.), *Toward a Science Campus in Milan*,
https://doi.org/10.1007/978-3-030-01629-6_14

14.1 Introduction

Since the early experiments of H. G. Dehmelt in the 1960s [1], Penning traps have very successfully lent themselves to a variety of uses in their capacity as devices for the confinement of electrically charged particles at low energy. The confinement, based on the superposition of an electrostatic well created by an array of suitably shaped electrodes and of a transverse-confining coaxial magnetic field, can be theoretically indefinite and is constrained only by mechanical and electrical imperfections or by collisions, which are usually limited by operation in ultra-high vacuum (UHV). This feature has allowed experimentalists to investigate single- or few-ion systems, measuring their fundamental atomic and nuclear properties or exploring their potential in quantum computing [2], as well as to store many-particle ensembles to observe their collective dynamics and equilibrium characteristics [3]. For the latter purpose cylindrical, elongated devices have been developed (Penning-Malmberg traps [4]), more apt to the confinement of large samples, whose behaviour is heavily influenced by the space-charge electric field, hence the name of nonneutral plasmas. In a typical regime where the plasma constituents are highly magnetized, i.e. their cyclotron motion has very small amplitude and period with respect to the other length and time scales of the system, the $\mathbf{E} \times \mathbf{B}$ -drift plasma dynamics is described by equations that are isomorphic to the Euler equations for a two-dimensional (2D) incompressible and inviscid fluid. In this respect, a trapped nonneutral plasma is a perfect laboratory for the study of transport phenomena, 2D fluid turbulence, self-organization.

Perturbations, both unwanted or intentionally exerted, of the natural plasma evolution shed light on the dynamics of these systems, and more specifically about the instabilities and the attainable equilibrium states. Such information is especially valuable when long-term confinement and manipulation of precious particle samples are critical, for instance antimatter synthesis [5]. A significant example of strongly-perturbed system is represented by in-trap generation of the electron plasma by ionization of the residual gas and progressive accumulation of the negative (electron) component. As we will show, this is possible applying a rather weak radio-frequency (RF) excitation to one of the trap electrodes, but comes at a cost, i.e. a reduced control over some of the most important parameters of the trapped column, specifically the transverse particle distribution and density. On the other hand, plasmas produced via this generation mechanism exhibit new and interesting properties with respect to those trapped by controlled injection from external sources. Some of the features observed in the realization of this technique and in the resulting plasmas can be exploited to adapt common trapped-plasma manipulation techniques to the present situation, which, in turn, both helps in the understanding of the physical phenomena and in overcoming the drawbacks and limitations of the technique itself in terms of stable, repeatable plasma samples. In the following we will review some of the most significant observations drawn from the studies we have conducted on RF-generated nonneutral plasmas in recent years.

14.2 Trapping Devices

Two set-ups were used to perform the experimental observations to be described in the following. Both devices are conceptually similar since they were conceived as Penning-Malmberg traps aimed at the confinement of electron plasmas, but differ in the range of operation with respect to some of the most significant parameters, namely the trapping volume and the magnetic field intensity of the solenoid coil placed around the vacuum vessel to provide a uniform magnetic field along the longitudinal axis over the whole trapping region. A sketch of the inner structure of the traps is shown in Fig. 14.1. The ELTRAP device [6] consists of a stack of 12 cylindrical electrodes of inner radius 45 mm. The 10 innermost electrodes can be biased to electrostatic potentials in the ± 100 V range, yielding a maximum trapping length of about 1 m. The axial magnetic field B_z can reach up to 0.2 T. The smaller ELTRAPPINO device is made out of 9 electrodes with a radius of 22.5 mm for a maximum trapping length of about 220 mm and a bias voltage range of ± 200 V. The magnetic field intensity is ≤ 0.88 T. The two set-ups also share diagnostic features: Some of their electrodes are split into two or more azimuthal sectors, which can be used as electrostatic pick-ups to detect the transverse motion of the plasma column or to apply RF excitations to manipulate it; At one end of the stack, a phosphor screen set at a voltage of some kilovolts can collect the ejected electrons. The light emitted by the decay of excited electrons in the active layer, recorded by a synchronized external camera, reveals the axially-integrated transverse distribution of the electron column at ejection from the trap. Both traps can reach a vacuum level in the UHV range, i.e. $\leq 10^{-8}$ mbar, thus limiting collisions and hence diffusion effects to the time scale of

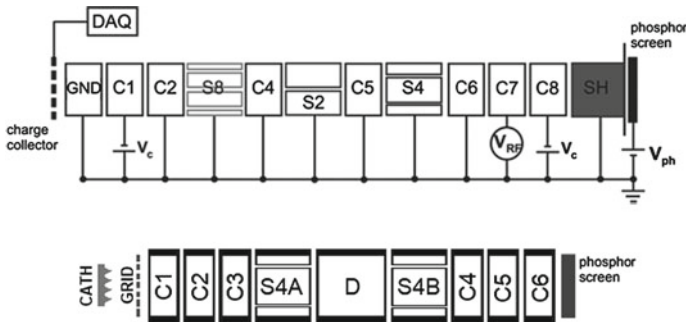


Fig. 14.1 Schematic outline of the ELTRAP and ELTRAPPINO electrode stacks. Top: ELTRAP stack, with a typical configuration for electron generation and trapping with electrodes C1 and C8 as endcaps at negative potentials $-V_c$ and the generation drive V_{RF} on C7. All other electrodes are grounded. Electrodes S2, S4, S8 are sectored into 2, 4, 8 azimuthal patches, respectively. A collector plate and a phosphor screen (set to a positive accelerating bias V_{ph}) are installed at the stack ends. Bottom: ELTRAPPINO electrode stack, with S4A and S4B four-fold split electrodes. The field-emission cathode and its electron extraction grid are sketched on the left, while a phosphor screen is placed at the opposite trap end

100 ms at least. The ELTRAPPINO device is also equipped with an external electron source made out of a matrix of field-emission tungsten tips and an acceleration grid. More details about the experimental set-ups are reported elsewhere [7, 8].

14.3 RF Generation

Conventional techniques to trap an electron plasma consist in the injection of a continuous or pulsed beam that is reflected by the endcap potential on the opposite side with respect to the injection one, and then trapped by raising the second endcap voltage. This requires a thermo-, photo- or field-emission cathode placed on the trap axis, thus blocking the line of sight, or adequate optical elements to steer on-axis the beam from an off-axis source. Depending on the type and geometry of the source, very high repeatability and flexibility in the transverse distribution can be obtained.

On the contrary, the way the electron plasma is produced in the experiments discussed here does not require a physical source apparatus. The trap is set in a confinement configuration, activating endcap potentials and magnetic field. A RF potential of the type $V_{\text{RF}} \cdot \sin(2\pi \nu_{\text{RF}} t)$ is applied to one of the inner electrodes. The few free electrons always present in the residual gas repeatedly interact with the periodically time-varying barrier represented by the localized RF drive in their longitudinal oscillation; over the course of many interactions, they undergo a stochastic heating that can reach up to the first ionization threshold of light gases and initiate a discharge. Similar chaotic heating models have been previously introduced in other contexts [9, 10]. Experimental observations demonstrate that RF drives with $V_{\text{RF}} = 1\text{--}10 V_{\text{pp}}$ and $\nu_{\text{RF}} = 1\text{--}30$ MHz are effective in the production of a confined electron column [7, 11]. The accumulation of electrons reaches detectable levels (charges of few picocoulombs, particle densities of the order of 10^5 cm^{-3}) in times of the order of some hundred milliseconds and within some seconds an equilibrium between continuous generation and losses is reached where the density profile is stable. Linear and nonlinear resonances of the RF field with three-dimensional single-particle or collective modes can also contribute to the heating effect and to the final states. These equilibrium states exhibit ample variations in transverse distribution, density and total electron charge; Fig. 14.2 shows an example of total charge and core density generated and trapped after 4.9 s of excitation at different frequencies in the 0.1–20 MHz range. Many of these plasmas are diffuse, quasi-axisymmetric and sometimes feature a hollow cross section. Due to the relatively low density and large shot-to-shot profile fluctuations, these are not of practical use for further experiments. Conversely, a notable phenomenon is the insurgence of high-density structures giving rise to coherent vortices from the diffuse background.

Figure 14.2 also highlights another feature of the RF generation process, i.e. a frequency threshold below which no confinement plasma is observed. As shown in the left diagram of Fig. 14.3, the threshold frequency ν_{thr} appears to depend linearly on the inverse of the trapping length L_{trap} , which confirms the heating to be mainly in the axial energy component [12] as the typical frequency range of applicability starts

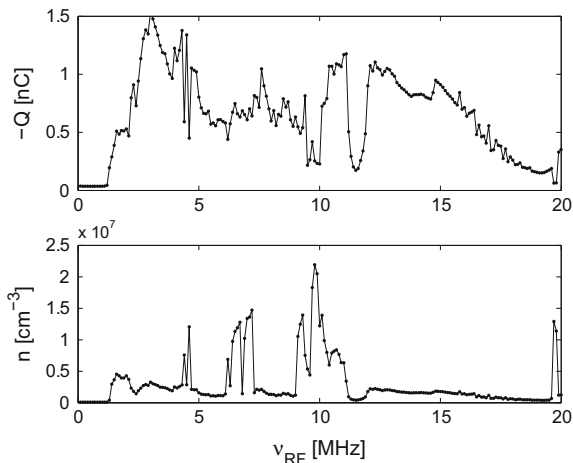


Fig. 14.2 Dependence of electron plasma charge and density on the RF generation frequency. Each data point represents the total confined charge (top) and core density (bottom) at ejection and dump onto the phosphor screen after 4.9 s of RF excitation. Trapping length 570 mm, magnetic field intensity 0.1 T, RF amplitude 1.5 V applied to a non-sectored electrode adjacent to an endcap, pressure about $1 \cdot 10^{-8}$ mbar

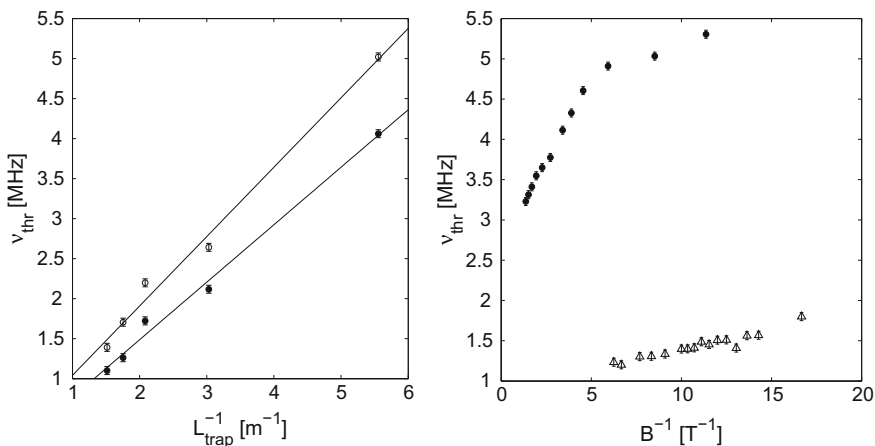


Fig. 14.3 Trends of threshold frequency for plasma generation. Left: dependence on the trapping length. The two data sets come from distinct experiments at slightly different pressures. Open circles: $p \simeq 0.9-1.2 \cdot 10^{-8}$ mbar. Full circles: $p \simeq 1.5-2.0 \cdot 10^{-8}$ mbar. Both experiments were performed in ELTRAP with a trapping length $L_{\text{trap}} = 570$ mm, a magnetic field of 0.1 T, and RF generation drive amplitude 1.5 V. The solid lines are linear fits of each data set against the inverse of the trapping length. Right: dependence on the magnetic field intensity. The two data sets shown come from two experiments performed in the two different traps. Open triangles: ELTRAP data, $L_{\text{trap}} = 570$ mm, RF generation drive amplitude 1.5 V, $p \simeq 1.5-2.0 \cdot 10^{-8}$ mbar. Full circles: ELTRAP-PINO data, $L_{\text{trap}} = 130$ mm, RF generation drive amplitude 3.5 V, $p \simeq 2.6-4.4 \cdot 10^{-7}$ mbar. See [7] for a discussion of uncertainties and related error bars (± 50 kHz)

at 1–5 MHz, corresponding to axial bounce frequencies (also proportional to L_{trap}^{-1}) for kinetic energies in the eV range. In the experimented range of magnetic fields, where electrons can always be considered to be highly magnetized, a monotonic dependence on the inverse of B_z is also observed (see right diagram of Fig. 14.2), although a clear understanding of the trend (which does not follow the usual B_z^{-2} scaling of classical diffusion) is still lacking.

14.4 Ion Trapping

A Penning trap can in principle only confine particles with one sign of charge. Nevertheless, as a dense electron column produces a negative space-charge potential, a configuration with grounded electrodes outside the endcaps may be sufficient to trap low-energy ions in the axial region including the endcaps themselves. This arrangement constitutes a special case of the more general *nested-trap* configuration, where opposite biasing of pair of electrodes is used to create a double-well potential shape. In particular, ions will mostly fall and accumulate in the endcap regions, but some of them may still repeatedly cross the entire trap. In both cases, the overall plasma confinement will be affected if at any time the amount of positive charge either stored in the endcaps or bouncing through the electron confinement length is significant. For instance, it has been demonstrated that even a very small degree of neutralization (ion-to-electron number ratios N_i/N_e of the order of 10^{-5}) may lead to destructive effects (trapped-ion diocotron instability) by radial drift and loss of the electron column [13].

In a series of tailored experiments we could demonstrate and quantitatively assess the trapping of an ion fraction [14, 15]. ELTRAP's C4 and C8 electrodes were used as endcaps and a RF drive was applied on C7 for a few seconds, until a steady state was reached. The electron plasma was ejected towards the phosphor screen setting C8 to ground. Electrode C4 was grounded after 4 ms in order to fully decouple the particle signals. Each grounding was accompanied by an RC discharge signal from the collector due to the ejection of ions trapped in the endcap regions. Several measurements with a range of trapping parameters (RF and magnetic fields) yielded electron samples of the order of 0.1–1 nC and positive charge depositions on the collector equivalent to fractions $N_i/N_e \simeq 10^{-2} - 10^{-1}$. A tailored experiment was performed to maximize and better resolve the ion storage properties. In this case one of the endcap regions was axially elongated by negatively biasing electrodes C2–C4 and the ion accumulation rate was measured repeating the generation and dump cycle for increasing confinement times. The result is summarized in Fig. 14.4, where the ion charge measured on the collector is plotted versus the confinement and excitation time. As the electron plasma reaches the equilibrium configuration (total charge about 800 pC), the trapped ion number also saturates to a fraction $N_i/N_e \simeq 5 \cdot 10^{-2}$ with an accumulation time scale of about 300 ms. The ion fractions result orders of magnitude larger than those observed in previous experiments [13], because of the long-time application of the RF drive, whose consequences on the plasma dynamics

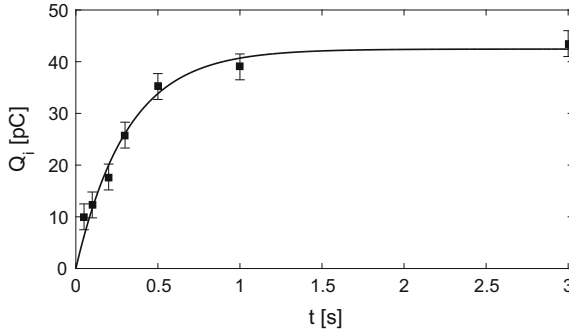


Fig. 14.4 Ion accumulation in the endcap region. The total ion charge Q_i accumulated in the negatively biased endcap is shown for increasing time of application of the ionizing RF signal (sinusoidal drive with $V_{RF} = 4.5$ V_{pp} and $\nu_{RF} = 10$ MHz on C7 electrode; $B = 0.12$ T). A fitting curve $\tilde{Q}[1 - \exp(-t/\tau)]$, $\tilde{Q} = 42.44$ pC, $\tau = 314$ ms is also plotted (solid line)

and equilibrium cannot be ignored. Indeed, if the RF generation drive is turned off an $l = 1$ diocotron instability can take place with growth time scales of 0.1–10 s, confirming the presence of ions in the electron trapping volume. Measurements reveal that the instability can be damped over times of some hundred milliseconds, as the electrons cool off and residual ions in the trapping volume are lost. The stability of an electron column despite the large ion fraction while the generation drive is active is a complicated phenomenon resulting from the balance between diocotron instabilities, particle losses and the continuous injection of charge, energy and angular momentum by the RF drive [14].

14.5 Multipolar Excitation and Compression

Diocotron waves are a class of density perturbations of the cross section of a non-neutral plasma column. Expressing the transverse density distribution $n(r, \vartheta)$ as a Fourier expansion in the azimuthal direction, the l -th wave mode will have the form $\delta n_l(r) \exp\{i[l\vartheta - \omega_l t]\}$, where δn_l is the amplitude and ω_l the oscillation frequency of the mode. For a thorough treatment of diocotron waves and their stability, see [16]. It is sufficient to remind here that in the linear regime, the frequency of the l -th mode for an initially circular, uniform-density distribution of radius R_p is $\omega_l = \omega_D \left[l - 1 + (R_p/R_W)^{2l} \right]$, with $\omega_D = ne/2\epsilon_0 B$ the rotation frequency of the plasma column, also called fundamental diocotron frequency. All modes are theoretically stable for a flat density profile; the most straightforward way to excite the l -th mode is by applying an external drive, oscillating at the mode frequency ω_l , to an azimuthally-split electrode in such a way that a $2l$ -polar boundary condition is produced. The multipolar drive is usually a small perturba-

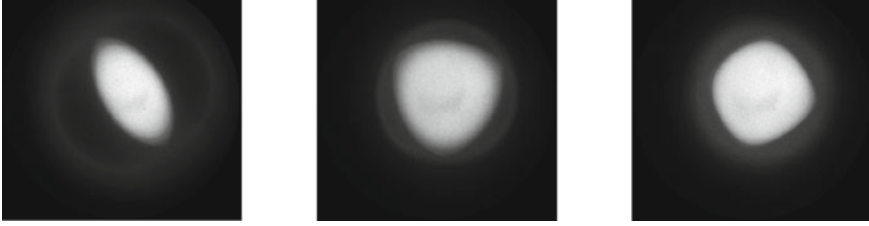


Fig. 14.5 Multipolar deformation of the plasma column upon resonant excitation of diocotron modes. From left to right, nonlinear diocotron modes of order 2, 3, 4 are excited by the application of periodic oscillating drives sweeping through the corresponding resonant mode frequency (17–18, 34–35 and 52–53 kHz sweeps, respectively). The drive amplitudes are 375 mV, 750 mV and 1 V, respectively. In order to excite a mode number l , a $2l$ -polar excitation is applied to a sectored electrode. The light intensity map (ascending from black to white) is normalized to a maximum density value of $1.2 \cdot 10^6 \text{ cm}^{-3}$

tion compared to the plasma potential; if the amplitude is too large, part of the plasma may be radially lost as the separatrix between closed and open equipotential lines is pushed towards the axis. As an example, an electron column with a flat density profile of about $1 \cdot 10^6 \text{ cm}^{-3}$, normalized radius $R_p/R_W \simeq 0.63$ and negligible radial offset was produced in ELTRAP applying a drive with $V_{\text{RF}} = 5.5 V_{\text{pp}}$, $\nu_{\text{RF}} = 7.5 \text{ MHz}$ on electrode C7 while biasing C2 and C8 at -80 V in a magnetic field of 0.12 T, reaching a steady state in few seconds. After turning off the generation drive, quadrupole, exapole and octupole excitations were applied for 400 ms on sectored electrodes, resulting in $l = 2$ (elliptic), $l = 3$ (triangular) and $l = 4$ (square) deformations of the circular cross section when the drive frequencies were close to the resonant frequency of the respective mode (Fig. 14.5). A broadening and upshift of the resonant frequency was found experimentally and attributed to a combination of deviations from the ideal step-like density profile, shot-to-shot profile fluctuations and increase in density of the highly deformed plasma core.

A very different situation occurs if the multipolar excitation is applied while the (higher-frequency) generation drive is still active. In this case, with the increase of the multipolar drive amplitude (which had destructive results when applied to a freely-evolving plasma), while the cross section remained approximately axisymmetric and centered on the longitudinal axis, a corresponding decrease in the electron charge was detected. This reduction, due to a concurrent shrinking of the cross section's mean radius, was nevertheless accompanied by a sensible increase in density of the remaining electron plasma [15]. Figure 14.6 shows the case of a quadrupolar drive, where the peak density increases up to a factor larger than 4. Preliminary measurements indicate that the density growth is maximum when the quadrupolar perturbation matches the $l = 2$ diocotron mode frequency. Similar density enhancement effects were found in the case of $l = 3$ and $l = 4$ excitations. This effect can hence be exploited to the aim of generating dense and centered vortices of practical use in further experiments requesting specific and well-controlled properties, thus

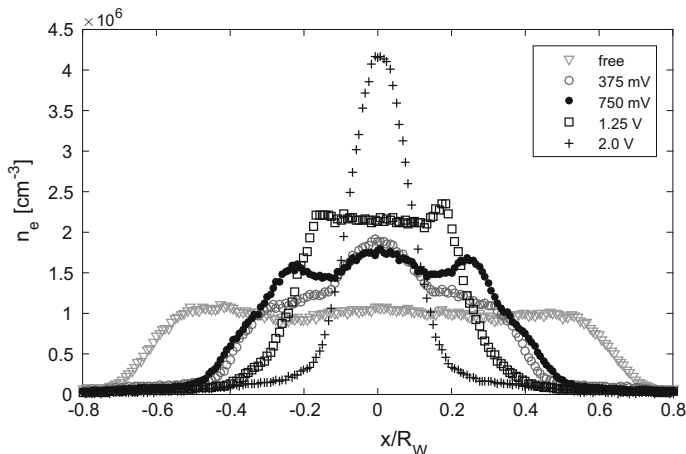


Fig. 14.6 Column compression by two-frequency excitation. The original plasma distribution (marked as ‘free’ in the legend), obtained with a 7.5 MHz, 1.5 V RF excitation applied to an azimuthally-symmetric electrode, shows a progressive reduction in radius and increase in peak density as a 400-ms quadrupolar excitation of increasing amplitude at the $l = 2$ diocotron mode frequency is added. The density profiles are plotted versus a transverse coordinate, normalized to the trap radius R_W . Quadrupolar drive amplitudes are listed in the legend

counteracting the inherent limitations of the RF-driven plasma generation technique, i.e. the lack of control and repeatability in the charge, position and density profile of vortex configurations.

14.6 Conclusions

We have reviewed some features of an original nonneutral plasma in-trap generation and confinement technique in a Penning-Malmberg trap exploiting a small-amplitude RF drive. The RF-based production technique presents apparent limitations, in that empirical adjustment of the trapping and excitation parameters is required to obtain plasmas with the desired tailored properties (global electron charge, repeatability, density profile), but also exhibits peculiar phenomena inherent to the specific process, which can be advantageously exploited. Regarding the largely predominant electron component, diffuse plasmas are more frequently observed. While they offer valuable insight into the early stages of plasma formation and accumulation, their low density, often accompanied by density profile and charge fluctuations, makes them difficult to study and manipulate. More interestingly, narrow-size, long-lived vortices can also be produced. A peculiarity of these structures is a very robust off-axis steady state. This comes out as a balance between the diocotron instability and the RF perturbation which counteracts the radial drift by feeding charge, energy and, depending on the azimuthal structure of the drive electrode, angular momentum.

The alteration of the natural equilibrium in the presence of the RF drive was further investigated by applying conventional manipulation techniques, yielding unexpected results. Specifically, the simultaneous use of the high-frequency, ionizing drive and of a multipolar perturbation at the associated diocotron mode frequency resulted in the formation of a more compact, yet denser and centered vortex. This is a particularly important result also regarding the control over the fundamental features of the electron plasmas, and in turn a key factor in improving both the understanding and the usability of RF-generated plasmas. For instance, specific shapes of the transverse distribution are required to test theoretical and numerical findings about the turbulent relaxation to equilibrium [17, 18].

We have also demonstrated the occurrence of nested trapping of an appreciable fraction of positive ions. Endcap potential modification by the accumulation of ions as well as repeated ion crossing of the electron trapping volume significantly contribute to modify the dynamics and equilibrium of the electron component. On the other hand, the first quantitative measurements of total ion charge and ion confinement time indicate opportunities of accumulation and manipulation of ion bunches and beamlets generated by means of this original production technique.

Acknowledgements The studies presented in this paper benefited from the financial contributions received in the framework of the projects ‘COOLBEAM’ and ‘PLASMA4BEAM’ (INFN Group V) and through the UNIMI funding Linea 2-Azione A 2014–2016.

References

1. H.G. Dehmelt, F.L. Walls, *Phys. Rev. Lett.* **21**, 127 (1968)
2. J. DiSciacca, M. Marshall, G. Gabrielse, S. Ettenauer, E. Tardiff, R. Kalra, D.W. Fitzakerley, M.C. George, E.A. Hessels, C.H. Storry, M. Weel, D. Grzonka, W. Oelert, T. Seifick, *Phys. Rev. Lett.* **110**, 130801 (2013)
3. G. Bettega, R. Pozzoli, M. Romé, *New J. Phys.* **11**, 053006 (2009)
4. J.H. Malmberg, J.S. deGrassie, *Phys. Rev. Lett.* **35**, 577 (1975)
5. G.B. Andresen, M.D. Ashkezari, M. Baquero-Ruiz, W. Bertsche, P.D. Bowe, E. Butler, C.L. Cesar, M. Charlton, A. Deller, S. Eriksson, J. Fajans, T. Friesen, M.C. Fujiwara, D.R. Gill, A. Gutierrez, J.S. Hangst, W.N. Hardy, R.S. Hayano, M.E. Hayden, A.J. Humphries, R. Hydromako, S. Jonsell, S.L. Kemp, L. Kurchaninov, N. Madsen, S. Menary, P. Nolan, K. Olchanski, A. Olin, A. Pusa, C.O. Rasmussen, F. Robicheaux, E. Sarid, D.M. Silveira, C. So, J.W. Storey, R.I. Thomson, D.P. van der Werf, J.S. Wurtele, Y. Yamazaki, *Nature Phys.* **7**, 558 (2011)
6. M. Amoretti, G. Bettega, F. Cavaliere, M. Cavenago, F. De Luca, R. Pozzoli, M. Romé, *Rev. Sci. Instrum.* **74**(9), 3991 (2003)
7. G. Maero, S. Chen, R. Pozzoli, M. Romé, *J. Plasma Phys.* **81**, 495810503 (2015)
8. M. Maggiore, M. Cavenago, M. Comunian, F. Chirulotto, A. Galatà, M. De Lazzari, A.M. Porcellato, C. Roncolato, S. Stark, A. Caruso, A. Longhitano, F. Cavaliere, G. Maero, B. Paroli, R. Pozzoli, M. Romé, *Rev. Sci. Instrum.* **85**, 02B909 (2014)
9. M.A. Lieberman, V.A. Godyak, *I.E.E.E. Trans. Plasma Sci.* **26**, 955 (1998)
10. J.L. Mateos, *Phys. Lett. A* **256**, 113 (1999)
11. B. Paroli, F. De Luca, G. Maero, R. Pozzoli, M. Romé, *Plasma Sources Sci. Technol.* **19**, 045013 (2010)

12. G. Maero, R. Pozzoli, M. Romé, S. Chen, M. Ikram, *JINST* **11**, C09007 (2016)
13. A.A. Kabantsev, C.F. Driscoll, *Fus. Sci. Technol.* **51**, 96 (2007)
14. G. Maero, *Il Nuovo Cimento C* **40**, 90 (2017)
15. G. Maero, S. Chen, R. Pozzoli, M. Romé, *A.I.P. Conf. Proc.* **1928**, 020009 (2018)
16. R.C. Davidson, *An Introduction to the Physics of Nonneutral Plasmas.* (Addison-Wesley Publishing Company, 1990)
17. S. Chen, G. Maero, M. Romé, *J. Plasma Phys.* **81**, 495810511 (2015)
18. M. Romé, S. Chen, G. Maero, *Plasma Phys. Control Fusion* **59**, 014036 (2017)

Chapter 15

Feeding Genetic Heterogeneity via a Smart Mutation Operator in the Memetic Phase Retrieval Approach



Marta Mauri, Davide Emilio Galli and Alessandro Colombo

Abstract A memetic algorithm is a stochastic optimization method obtained by hybridizing an evolutionary approach with common deterministic optimization procedures. The recently introduced Memetic Phase Retrieval (MPR) approach exploits this synergy to face the so-called *phase retrieval problem* in Coherent Diffraction Imaging (CDI). Here we focus on the development of a *smart mutation* genetic operator; our aim is the improvement of MPR performance by continually feeding with *relevant* information the genetic heritage of the population of candidate solutions. Remarkably, statistical tests on synthetic CDI data performed using MPR enhanced via a smart mutation operator reveal a smaller reconstruction error with respect to an MPR implementation supplied with a blind random mutation only.

15.1 Introduction

The phase problem consists in the loss of information affecting the recovery of a function, $\rho(\mathbf{x})$, when only the modulus of its Fourier transform, $|\hat{\rho}(\mathbf{k})| = |\mathcal{F}[\rho(\mathbf{x})](\mathbf{k})|$, is known. This issue is present in many different fields, from crystallography [1] to astronomy [2], including electron microscopy [3] and optical imaging [4]. In particular, the phase problem affects the technique known as Coherent Diffraction Imaging (CDI) [5], which is at the basis of an increasingly important field of research, currently emerging thanks to the development of a new generation of coherent light sources such as X-Ray Free Electron Lasers [6].

CDI is a lens-less technique aimed at the quantitative imaging of matter which makes use of strongly coherent radiation, usually photons or electrons, impinging on a sample. CDI is not affected by lens aberrations and the achievable resolution depends, in principle, only on the radiation wavelength, allowing a quantitative imaging of matter down to the sub-atomic scale. In a CDI experiment, the diffraction pattern $I(\mathbf{k})$ is acquired in *far field* conditions by an optical detector and turns out to be equal, up to a normalization factor, to the square modulus of the Fourier transform of the sample

M. Mauri · D. E. Galli · A. Colombo (✉)
Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Milan, Italy
e-mail: alessandro.colombo6@unimi.it

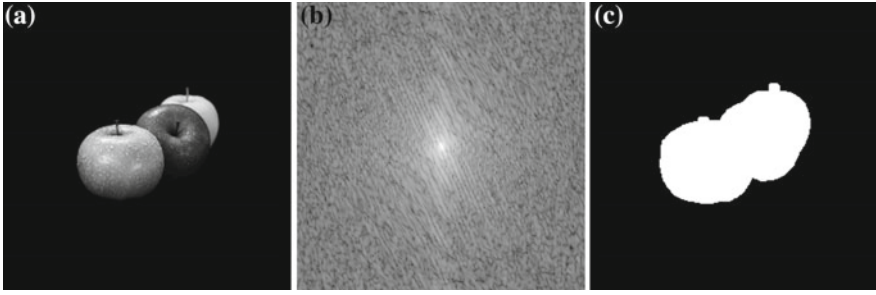


Fig. 15.1 Example of CDI data: (panel **a**) target density $\rho(\mathbf{x})$, to be retrieved from the knowledge of the diffraction pattern $I(\mathbf{k}) = |\tilde{\rho}(\mathbf{k})|^2$ ((panel **b**), shown in logscale, and (panel **c**) the support function $S(\mathbf{x})$. (panel **b**) and (panel **c**) will be exploited to test the algorithm in the following sections

spatial density, $I(\mathbf{k}) = |\tilde{\rho}(\mathbf{k})|^2$. Usually, CDI detectors are optical devices, such as Charge-Coupled Devices (CCD), able to acquire only the intensity of the impinging radiation. Accordingly, the phase information, $\arg[\tilde{\rho}(\mathbf{k})]$, is completely lost in the measurement process, turning the retrieval of the scattering function $\rho(\mathbf{x})$ into a complex task. Since the reciprocal-space experimental constraint $I(\mathbf{k}) = |\tilde{\rho}(\mathbf{k})|^2$ does not ensure the existence of a unique solution, further information is required. This information is provided by a real-space constraint encoded in a binary function $S(\mathbf{x})$, called *support function*, such that $S(\mathbf{x}) = 1$ where $\rho(\mathbf{x}) \neq 0$, and $S(\mathbf{x}) = 0$ elsewhere. Given suitable conditions, according to the Shannon theorem [1], the density $\rho(\mathbf{x})$ which fulfills both the experimental constraint and the support constraint is unique and represents the solution to the phase problem. Figure 15.1a shows an example of a density to be recovered, given its diffraction pattern, Fig. 15.1b, and using the support function depicted in Fig. 15.1c.

In order to face the phase retrieval problem, several deterministic iterative procedures have been developed so far [7]; among these, the Error Reduction (ER) and the Hybrid Input Output (HIO) [8] were the first effective algorithms for CDI data analysis. They exploit the cyclical imposition of the experimental and the support constraints in order to converge towards the solution $\rho(\mathbf{x})$. Error Reduction is the simplest method and it acts as follows:

$$\rho_{n+1}(\mathbf{x}) = S(\mathbf{x}) \cdot \mathcal{F}^{-1}[\sqrt{I(\mathbf{k})} \exp(i\Phi_n(\mathbf{k}))](\mathbf{x}), \quad (15.1)$$

where $\Phi_n(\mathbf{k}) = \arg[\tilde{\rho}_n(\mathbf{k})]$. Going back and forth from the reciprocal to the real space, these methods attempt to minimize the following error functional:

$$E[\rho_n] = \sum_{\mathbf{k}} |\sqrt{I(\mathbf{k})} - |\tilde{\rho}_n(\mathbf{k})||^2 \quad (15.2)$$

turning the phase retrieval problem into an optimization task. In particular, the Error Reduction algorithm represents a steepest-descent optimization method for the

quantity defined in (15.2) [8]. The retrieved density is assumed to be the one which minimizes $E[\cdot]$. The ideal solution to the phase problem is attained when $E[\rho_n] = 0$, while, in presence of experimental noise, the minimum reachable error is higher.

The number of independent parameters involved in the optimization procedure depends on the resolution of the acquired diffraction pattern $I(\mathbf{k})$: typical sizes are around 1024×1024 pixels, i.e. about 10^6 intensity values. The real-space scattering function $\rho(\mathbf{x})$ is stored in a matrix with the same dimensions, but clearly $\rho(\mathbf{x})$ has to be retrieved only where $S(\mathbf{x}) = 1$. Usually, the support function extension is few tenths of the whole matrix, so that about 10^5 unknowns have to be retrieved. This huge number of involved parameters makes the phase retrieval problem a really complex optimization task; moreover, lack of information about the experimental diffraction pattern and the support function makes the problem even harder.

Since deterministic algorithms treat the phase retrieval problem as an optimization task, the main issue concerning their usage is the stagnation in local minima, preventing them from finding the solution [9]. A common way to overcome stagnation is to accomplish many independent phase retrieval procedures from different starting points $\{\rho_0^{(j)}(\mathbf{x})\}$ and then selecting the ones with lower error, thus performing a *random search* optimization. Other possibilities employ more sophisticated heuristic optimization methods, as in the recently introduced Memetic Phase Retrieval approach [10].

15.2 Memetic Phase Retrieval Approach

Genetic Algorithms (GAs) [11] represent a common heuristic strategy to deal with optimization problems. GAs imitate the *survival-to-fitness* process of natural selection acting on a population of candidate solutions. Each individual in this population is uniquely defined by its genetic pool, i.e. a set of values for the parameters involved in the optimization problem. GAs perform an evolutionary process which affects these genetic pools in order to optimize a specific *fitness* function. The presence of a population, along with an embedded selection principle, reduces the probability of getting stuck in local optima, increasing the algorithm efficiency. Nevertheless, in spite of their computational intelligence, GAs are not suitable to face optimization problems where the number of parameters is very large and the evaluation of the fitness function is computationally demanding. This is exactly the CDI case, where the number of parameters usually is about 10^5 and the evaluation of the fitness function involves the Fourier transform. For these reasons, our heuristic approach to the phase problem exploits a Memetic Algorithm (MA) [12] obtained by hybridizing the evolutionary method with usual deterministic-iterative algorithms. The Memetic Phase Retrieval (MPR) approach [10] takes advantage of the synergy of both optimization strategies allowing a more accurate reconstruction of the unknown density $\rho(\mathbf{x})$. MPR has been recently applied to the phase retrieval of electron diffraction data leading to excellent results [10].

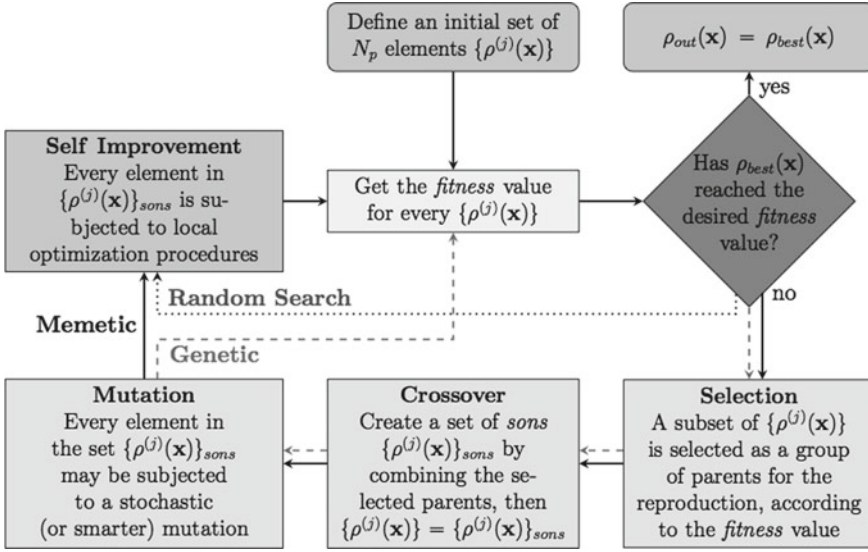


Fig. 15.2 Flowchart of the heuristic phase retrieval algorithms discussed in the text. Three paths are shown, corresponding to random search, genetic algorithm and memetic phase retrieval approaches

As in the case of standard GAs, our approach manipulates genetic information exploiting three typical genetic operators, namely Selection, Crossover and Mutation. Additionally, in MPR each individual undergoes a procedure of *self-improvement* via deterministic algorithms (see Fig. 15.2).

The goal of the memetic procedure is to maximize the fitness function $E[\cdot]^{-1}$, namely minimize the image reconstruction error. Starting from an initial population $\{\rho_0^{(j)}(\mathbf{x})\}$ of N_p elements, MPR creates subsequent generations of new individuals: according to the fitness value, some elements are selected as a group of parents for the reproduction and then combined to produce a new population, which replaces the previous one. Once $\{\rho^{(j)}(\mathbf{x})\}$ has been sorted by decreasing fitness values, parents for crossover are selected by choosing individuals with lower index s , i.e. higher fitness. The index is extracted via a *rigged roulette*, which turns a flat random distribution into an unbalanced one thanks to the parameter $r \geq 1$ through the following relation:

$$s = \lfloor \{\text{rand}[0, 1)\}^r \cdot N_p \rfloor + 1. \quad (15.3)$$

The Crossover operator mixes the parents' genetic heritage to create a new population. We used a method known as differential crossover [13] that selects and combines four parents, with indexes s_1, s_2, s_3, s_4 , to create a new individual $\tilde{\rho}_{\text{son}}(\mathbf{k})$. It acts as follows, depending on the differential coefficient D_c and a balancing coefficient C :

$$\tilde{\rho}_{\text{son}}(\mathbf{k}) = \begin{cases} \tilde{\rho}^{(s_1)}(\mathbf{k}) & \text{if rand } [0, 1) > C \\ \tilde{\rho}^{(s_2)}(\mathbf{k}) + D_c \cdot [\tilde{\rho}^{(s_3)}(\mathbf{k}) - \tilde{\rho}^{(s_4)}(\mathbf{k})] & \text{otherwise.} \end{cases} \quad (15.4)$$

The operations designed to *feed* the individual's genetic pool are represented by the Self Improvement, the Crossover and the Mutation operators. Within the original MPR implementation, the Mutation operator was a stochastic modification of some individuals' genetic pool $\rho^{(j)}(\mathbf{x})$ obtained by randomly changing the respective phases $\Phi^{(j)}(\mathbf{k}) = \arg[\tilde{\rho}^{(j)}(\mathbf{k})]$. This mutation can be labelled as *blind mutation*, because every parameter of the optimization problem has the same chance to run into a mutation. In the present work, our goal is to provide a better *feeding* for the genetic heritage of the evolving population, through the introduction of a *smart mutation* operator.

15.3 Development of a Smart Mutation Operator

Within the MPR approach, the action of a mutation operator is especially useful when the optimization is stuck in a local optimum due to a too much similar genetic heritage among the individuals, so that neither the Crossover nor the deterministic algorithms are able to escape this convergence basin. In such a situation, a mutation can provide a *smart feeding* of the genetic heritage by focusing its altering action where the candidate solutions' phases are more likely to encode inappropriate information. Moreover, it would be more efficient to focus the *smart mutation* on the parameters that are not involved in the action of the deterministic algorithms, in order not to interfere with them.

First of all, let's try to approximately identify some regions in the reciprocal space where, due to the presence of a reconstruction error, it is most likely that a mutation of the phases would feed with relevant information the evolutionary process. The only a-priori knowledge about the solution $\rho(\mathbf{x})$ of the phase retrieval problem is the modulus of its Fourier Transform $|\tilde{\rho}(\mathbf{k})| = \sqrt{I(\mathbf{k})}$, i.e. the square root of the experimental data. This implies that, given an estimation of the solution $\rho'(\mathbf{x})$ satisfying the support constraint, we are able to evaluate the discrepancy between the modulus of its Fourier Transform and the experimental data, that is $D(\mathbf{k}) = ||\tilde{\rho}(\mathbf{k})| - |\tilde{\rho}'(\mathbf{k})||$. Thanks to the fact that the solution $\rho(\mathbf{x})$ is a fixed point for (15.1), and exploiting the convolution theorem, we can write:

$$\tilde{\rho}(\mathbf{k}) = \mathcal{F}[S(\mathbf{x})] * [\sqrt{I(\mathbf{k})}e^{i\Phi(\mathbf{k})}]. \quad (15.5)$$

On the other hand, our estimation $\rho'(\mathbf{x})$ derives from a wrong retrieval of the phases, which differ from the true phases by a quantity $\Delta\Phi(\mathbf{k})$, so that we can write:

$$\tilde{\rho}'(\mathbf{k}) = \mathcal{F}[S(\mathbf{x})] * [\tilde{\rho}(\mathbf{k})e^{i\Delta\Phi(\mathbf{k})}]. \quad (15.6)$$

Exploiting the linearity of the convolution we obtain:

$$\tilde{\rho}(\mathbf{k}) - \tilde{\rho}'(\mathbf{k}) = \mathcal{F}[S(\mathbf{x})] * [\tilde{\rho}(\mathbf{k})(1 - e^{i\Delta\Phi(\mathbf{k})})]. \quad (15.7)$$

At this point, we introduce a convenient approximation: we assume that the convolution with $\mathcal{F}[S(\mathbf{x})]$ is negligible. In real applications, where the area filled by the support function is considerable [1], approximating $\mathcal{F}[S(\mathbf{x})]$ to a delta function can be considered not so inadequate. Moreover, it's worth noting that our aim is simply to approximately identify favourable regions of the reciprocal space where the mutation operator should act. Since we assumed the convolution operation to be avoidable, (15.7) reads:

$$\tilde{\rho}(\mathbf{k}) - \tilde{\rho}'(\mathbf{k}) \simeq \tilde{\rho}(\mathbf{k})(1 - e^{i\Delta\Phi(\mathbf{k})}), \quad (15.8)$$

Exploiting the inequality $||z_1| - |z_2|| \leq |z_1 - z_2|$, with $z_1, z_2 \in \mathbb{C}$, and recalling the definition of $D(\mathbf{k})$, we obtain the relation:

$$D(\mathbf{k}) = ||\tilde{\rho}(\mathbf{k})| - |\tilde{\rho}'(\mathbf{k})|| \leq |\tilde{\rho}(\mathbf{k}) - \tilde{\rho}'(\mathbf{k})| \simeq \sqrt{I(\mathbf{k})}|1 - e^{i\Delta\Phi(\mathbf{k})}|, \quad (15.9)$$

which leads to the useful condition:

$$\Delta M(\mathbf{k}) \doteq \frac{D(\mathbf{k})}{\sqrt{I(\mathbf{k})}} \lesssim |1 - e^{i\Delta\Phi(\mathbf{k})}|, \quad (15.10)$$

where $\Delta M(\mathbf{k})$ is called *error map* and measures a relative modulus variation. We conclude that phases are more likely to be incorrect where the reconstructed moduli $|\tilde{\rho}'(\mathbf{k})|$ deviate from the experimental data $|\tilde{\rho}(\mathbf{k})|$. The analytical argument is strengthened by an empirical proof, shown in Fig. 15.3, which displays an imposed phase variation, $\Delta\Phi(\mathbf{k})$, on the Fourier transform of the density depicted in Fig. 15.1a and the induced error map, $\Delta M(\mathbf{k})$. The correlation between $\Delta\Phi(\mathbf{k})$ in Fig. 15.3a and $\Delta M(\mathbf{k})$ in Fig. 15.3b is particularly evident.

Since the mutation operator is aimed to overcome the stagnation problem, it is supposed to have an effect only where the deterministic self-improvement procedures are stuck and don't succeed in changing the phase values anymore. Such information can be extracted by evaluating the action of the Error Reduction algorithm between two subsequent iteration steps. This can be obtained making use of (15.1) to compute the difference $\Delta\tilde{\rho} = \tilde{\rho}_{n+1}(\mathbf{k}) - \tilde{\rho}_n(\mathbf{k})$, and then to evaluate

$$\Delta G(\mathbf{k}) = \frac{|\Delta\tilde{\rho}|}{\max_{\mathbf{k}} |\Delta\tilde{\rho}|}, \quad (15.11)$$

which ranges between 0 and 1. At this point we create a *mutation map* M_{mut} by combining $\Delta G(\mathbf{k})$ and $\Delta M(\mathbf{k})$ in the following way:

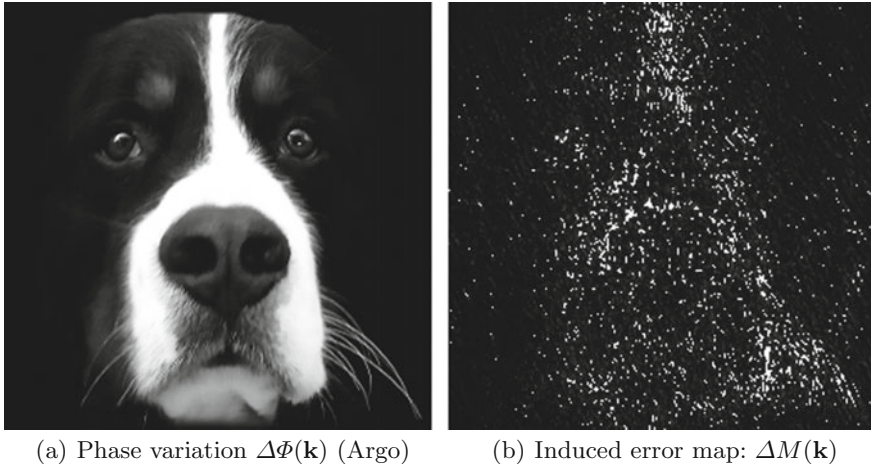


Fig. 15.3 Error map obtained via a phase variation for the Fourier transform of the density in Fig. 15.1a. Phase variation ranges from 0 (black) to π (white). Induced error map ranges from 0 (black) to ≥ 0.2 (white)

$$M_{\text{mut}}(\mathbf{k}) = \Delta M(\mathbf{k}) \cdot [1 - \Delta G(\mathbf{k})], \quad (15.12)$$

so that $M_{\text{mut}}(\mathbf{k}) \in [0, 1]$ represents a function of the reciprocal space which takes higher values where an error on the phase is expected and the iterative algorithm is not acting. Each individual in every generation undergoes the mutation process with a certain probability, governed by the parameter *mut_prob*, which ranges from 0 (no mutation at all) to 1 (every element undergoes a mutation). The smart mutation genetic operator modifies the individuals' genetic pool acting as follows:

$$\Phi_n^{(\text{mut})}(\mathbf{k}) = \Phi_n(\mathbf{k}) + M_{\text{mut}}(\mathbf{k})^{\text{mut_level}} \cdot \text{rand}[-\pi, \pi] \quad (15.13)$$

where the parameter *mut_level* represents the variation intensity. It's worth noting that the mutation operation depends strongly on the mutation map, including also a stochastic term which corresponds to a uniform random distribution with values among $[-\pi, \pi]$.

15.4 Tests and Results

We have tested the mutation operator performance on synthetic CDI data, thus considering a case where perfect knowledge of $I(\mathbf{k})$ is present, and we have evaluated its computational intelligence through the analysis of the reconstruction error defined in (15.2). The usage of synthetic data provides a clear evidence of the mutation effect without being affected by the presence of experimental noise. We carried out

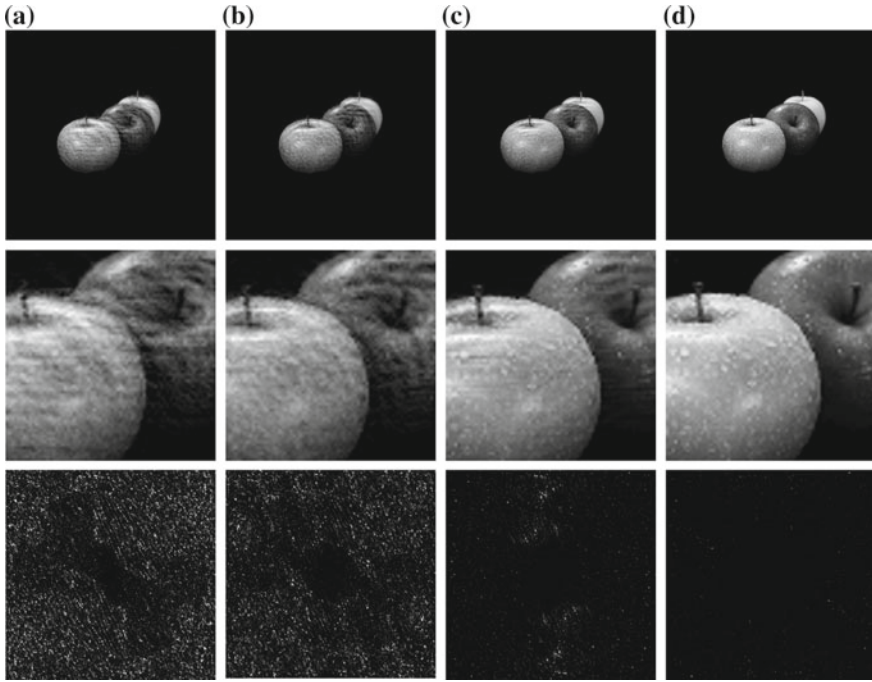


Fig. 15.4 The first row displays the best reconstructed densities, $\rho_{\text{best}}(\mathbf{x})$, obtained with (panel a) a random search, (panel b) MPR without mutation, (panel c) MPR with blind mutation, (panel d) MPR with smart mutation. In the second row, a detail of the reconstructed density is shown. The last row contains the corresponding error maps $\Delta M(\mathbf{k})$

simulations with populations composed of $N_p = 2176$ individuals, being each of them a 256×256 matrix, and running for about 500 generations. The input data for the phase retrieval procedure are shown in Fig. 15.1; the parameters *mut_level* and *mut_prob* have been set equal to 0.3 and 0.5, respectively. The Self Improvement step performed about 50 iterations for each generation.

We have compared MPR enhanced via the smart mutation operator with other three approaches: *random search*, MPR deprived of the mutation process and MPR implementation supplied with a *blind random mutation*. The first row of Fig. 15.4 shows the reconstructed densities, $\rho_{\text{best}}(\mathbf{x})$, by means of these different heuristic approaches, being ρ_{best} the individual with the highest fitness in the last generation. The second row of Fig. 15.4 shows a detail of $\rho_{\text{best}}(\mathbf{x})$ and the third row in the same figure displays the corresponding *error maps* $\Delta M(\mathbf{k})$. Remarkably, the quality of the reconstruction increases monotonically from A (random search) to D which corresponds to the MPR implementation supplied with the smart mutation operator. This is particularly evident by comparing the corresponding error maps in the third row; $\Delta M(\mathbf{k})$ goes to zero when the retrieved moduli agree with the experimental ones. The greyscale map used in Fig. 15.4 goes from black (totally correct: $\Delta M(\mathbf{k}) = 0$) to

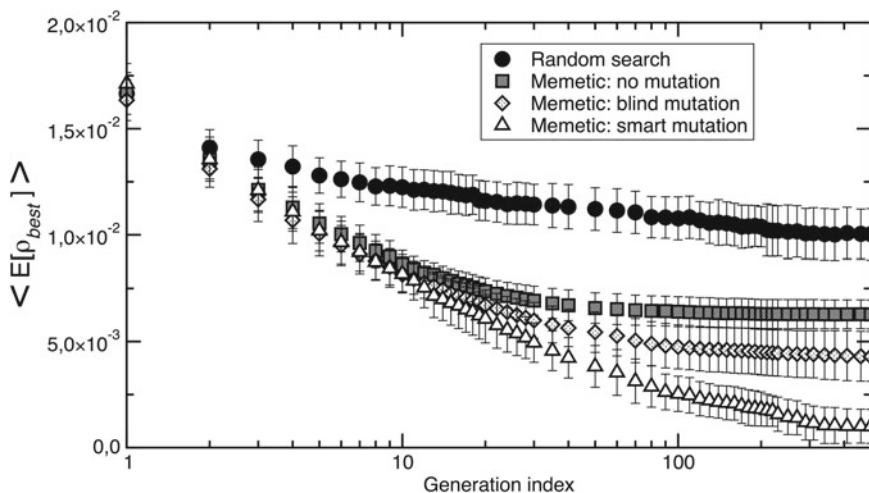


Fig. 15.5 Mean error value $\langle E[\rho_{\text{best}}] \rangle$ as a function of the generation index for different heuristic approaches as mentioned in the text. Error bars represent the width of the distributions for the obtained values of $E[\rho_{\text{best}}]$

white ($\Delta M(\mathbf{k}) \geq 1$). In the smart mutation case, it is evident that for most \mathbf{k} values this discrepancy is essentially zero.

In order to strengthen the evidence that MPR supplied with the smart mutation operator leads to better reconstructions even with respect to MPR supplied with a blind mutation, a statistical analysis of the heuristic approaches' performance is needed. We have carried out several independent phasing procedures with different random number generator seeds obtaining for each case an average value for the error functional, $\langle E[\rho_{\text{best}}] \rangle$, and the width of the distributions of its values, i.e. the standard deviations. Figure 15.5 displays $\langle E[\rho_{\text{best}}] \rangle$ as a function of the generation index. Note that the error bars depicted in Fig. 15.5 represent the standard deviations of the distributions of $E[\rho_{\text{best}}]$; one can appreciate that the different distributions are well separated. This statistical analysis thus confirms that the smart mutation implemented in MPR allows to systematically get a smaller $E[\rho_{\text{best}}]$, therefore turning out to be more efficient even with respect to MPR supplied with a blind random mutation. This is a remarkable result which shows that the implemented smart mutation operator is able to *feed* the genetic heterogeneity with relevant information improving convergence towards a better phase retrieval and image reconstruction.

In order to highlight the potentiality of our mutation operator, we have applied MPR to the reconstruction of a synthetic electron atomic potential (which is the equivalent of the density $\rho(\mathbf{x})$ in electron CDI) of a SrTiO₃ nanocrystal. Some defects like vacancies and misplaced atomic columns have been artificially inserted in a simulated electron diffraction data, whose dimension is 1024 × 1024 pixels (see Fig. 15.6a). In addition to the complications coming from the higher image resolution, the phasing procedure is made even more difficult due to the fact that the ideal crystal,

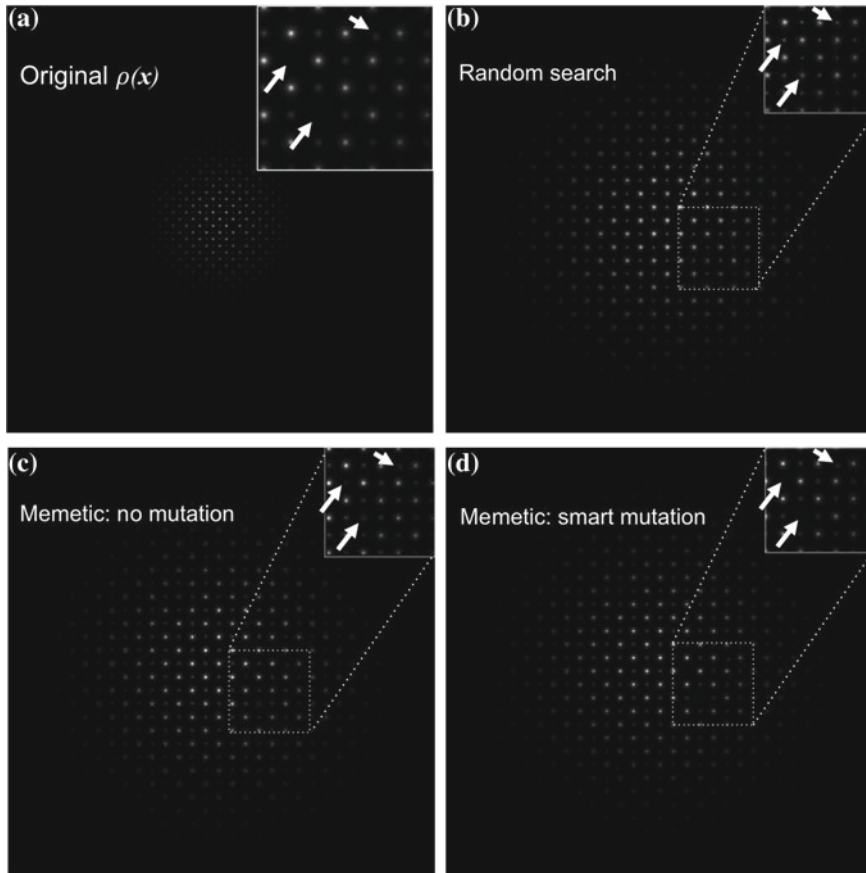


Fig. 15.6 The figure shows the original density, $\rho(\mathbf{x})$, to be recovered (panel **a**) and the obtained reconstructions for three different cases: random search case (panel **b**), memetic case without mutation (panel **c**) and memetic case with smart mutation (panel **d**). In the insets, a portion of the recovered densities $\rho_{\text{best}}(\mathbf{x})$ are enlarged in order to allow a better visualization of the retrieved defects

i.e. perfect lattice, is a local optimum located really close to the solution, making the stagnation phenomenon very likely. In Fig. 15.6 we compare the results obtained using a random search algorithm, MPR without mutation and MPR enhanced via the smart mutation. The crystal structure is recovered by all approaches, but it's worth noting that MPR enhanced via the smart mutation allows the best reconstruction of the crystal defects, as shown in the insets of Fig. 15.6.

15.5 Conclusions

This work represents a further development of the Memetic Phase Retrieval (MPR) approach, which is a memetic algorithm that exploits the synergy of evolutionary and deterministic optimization procedures to face the *phase retrieval* problem in the field of Coherent Diffraction Imaging. We have introduced a *smart mutation* genetic operator in order to get around possible stagnation issues affecting MPR. This operator manages to randomly vary only those parameters of the candidate solutions that are more likely to contain inappropriate information and that are not already altered by the action of the deterministic algorithms. A statistical analysis of quantitative results concerning the phase retrieval of synthetic data shows a systematic improvement in performance with respect also to MPR supplied only with a *blind mutation*. The introduction of this *smart mutation* in the Memetic Phase Retrieval approach turns out to be a relevant step towards a more accurate imaging of matter via coherent diffraction experiments, showing that computational intelligence can represent a valuable strategy in this research field.

Acknowledgements We acknowledge L. De Caro, E. Carlino and F. Scattarella for useful discussions. This work was supported by the NOXSS PRIN (2012Z3N9R9) project. We acknowledge the CINECA and Regione Lombardia LISA award LI05p-PUMAS, the CINECA ISCRA-C award IMAGES and the CINECA ISCRA-B award MEMETICO for the availability of high performance computing resources and support.

References

1. D. Sayre, Some implications of a theorem due to Shannon. *Acta Crystallogr.* **5**(6), 843 (1952)
2. C. Fienup, J. Dainty, Phase retrieval and image reconstruction for astronomy. in *Image Recovery: Theory and Application* (1987), pp. 231–275
3. L. De Caro, E. Carlino, F.A. Vittoria, D. Siliqi, C. Giannini, Keyhole electron diffractive imaging (KEDI). *Acta Crystallogr. Sect. A* **68**, 687–7026 (2012)
4. Y. Shechtman, Y.C. Eldar, O. Cohen, H.N. Chapman, J. Miao, M. Segev, Phase retrieval with application to optical imaging: a contemporary overview. *IEEE Signal Process. Mag.* **32**(3), 87–109 (2015)
5. L. De Caro, E. Carlino, D. Siliqi, C. Giannini, Coherent diffractive imaging: from nanometric down to picometric resolution, in *Handbook of Coherent-Domain Optical Methods* (Springer, 2013), pp. 291–314
6. M. Altarelli, R. Brinkmann, M. Chergui, W. Decking, B. Dobson, S. Düsterer, G. Grübel, W. Graeff, H. Graafsma, J. Hajdu et al., The European X-ray free-electron laser, in *Technical Design Report, DESY 97* (2006), pp. 1–26
7. S. Marchesini, Invited article: a unified evaluation of iterative projection algorithms for phase retrieval. *Rev. Sci. Instrum.* **78**(1), 011301 (2007)
8. J.R. Fienup, Phase retrieval algorithms: a comparison. *Appl. Opt.* **21**(15), 2758–2769 (1982)
9. J.R. Fienup, C.C. Wackerman, Phase-retrieval stagnation problems and solutions. *J. Opt. Soc. Am. A* **3**(11), 1897–1907 (1986)
10. A. Colombo, D.E. Galli, L. De Caro, F. Scattarella, E. Carlino, Facing the phase problem in coherent diffractive imaging via memetic algorithms. *Sci. Rep.* **7**(42236) (2017)
11. D.E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, 1st edn. (Addison-Wesley Longman Publishing Co. Inc., Boston, MA, USA, 1989). ISBN: 0201157675

12. P. Moscato et al., On evolution, search, optimization, genetic algorithms and martial arts: towards memetic algorithms, in *Caltech Concurrent Computation Program, C3P Report* (1989), p. 826
13. R. Storn, K. Price, Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces. *J. Glob. Optim.* **11**(4), 341–359 (1997)

Chapter 16

Jack on a Devil's Staircase



Andrea Di Gioacchino, Marco Gherardi, Luca Guido Molinari
and Pietro Rotondo

Abstract We review a simple mechanism for the formation of plateaux in the fractional quantum Hall effect. It arises from a map of the microscopic Hamiltonian in the thin torus limit to a lattice gas model, solved by Hubbard. The map suggests a Devil's staircase pattern, and explains the observed asymmetries in the widths. Each plateau is a new ground state of the system: a periodic Slater state in the thin torus limit. We provide the unitary operator that maps such limit states to the full, effective ground states with same filling fraction. These Jack polynomials generalise Laughlin's ansatz, and are exact eigenstates of the Laplace-Beltrami operator. Why are Jacks sitting on the Devil's staircase? This is yet an intriguing problem. Talk given in Milan, Congresso di Dipartimento 2017 (L.G.M.).

16.1 The Quantum Hall Effects

Since the discoveries of the integer [10] and of the fractional [21] quantum Hall effects, the phenomena have been an amazing source of inspiration for experimental and theoretical physics, with deep intersections with mathematics. After decades, the experimental results stand neat and beautiful (Fig. 16.1).

A gas of high mobility electrons at the interface of a heterostructure, refrigerated at mK temperature and in magnetic fields of one or several tesla, exhibits a quantisation of the resistance orthogonal to the current, in units of a fundamental constant, at integer or fractional values of ν :

A. Di Gioacchino · M. Gherardi · L. G. Molinari (✉)
Dipartimento di Fisica, Università degli Studi di Milano and I.N.F.N. sezione di Milano,
Via Celoria 16, 20133 Milan, Italy
e-mail: luca.molinari@unimi.it

P. Rotondo
University of Nottingham, Nottingham, UK

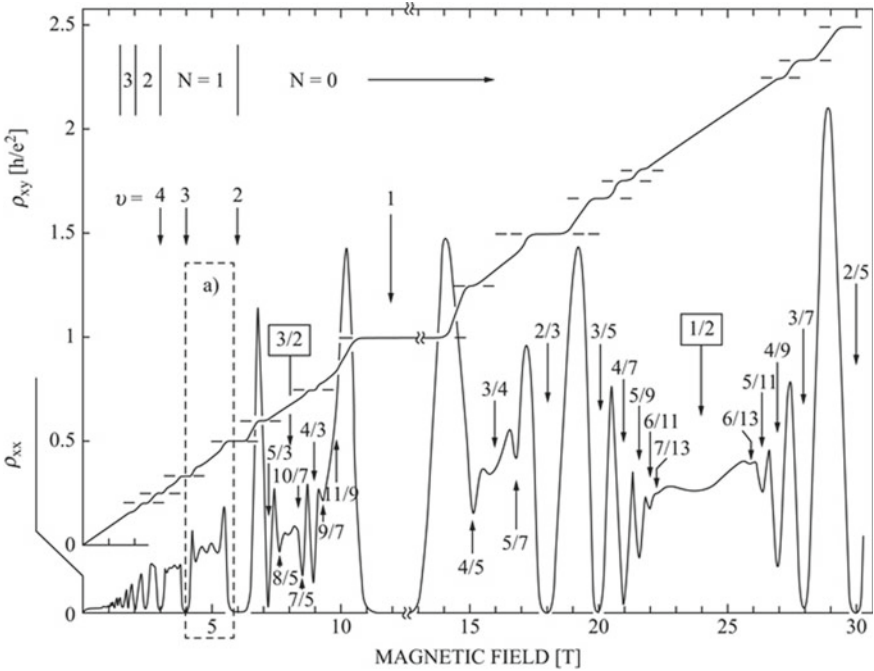


Fig. 16.1 The experimental ladder of plateaux for ρ_{xy} (in units of R_H) as a function of the magnetic field. Note the asymmetry in width of the plateaux ν and $1 - \nu$ (reprinted figure with permission, from [22]. Copyright 1987 by the American Physical Society)

$$R_{xy} = \left(\frac{h}{e^2} \right) \frac{1}{\nu} \tag{16.1}$$

The parameter is the filling fraction $\nu = n/g$, where n and $g = eB/hc$ are the number of electrons and the degeneracy of a Landau level per unit area. Substitution gives $R_{xy} = B/(enc)$: this is Hall’s law resulting from the balance of the Lorentz force with the electric force created by a charge gradient. However, in a quantum regime, the transverse resistance does not vary continuously with the field B : the linear slope is a staircase.

The integer plateaux $\nu = 1, 2, 3, \dots$ (IQHE) came first; next, by achieving higher magnetic fields, mobility, and lower temperatures, the prominent plateaux $\nu = 1/3$ and $2/3$ were discovered. In the years, 60 or more fractional values have been observed (FQHE), the values $\nu < 1$ corresponding to a partial filling of the lowest Landau level. While experiments are performed on micrometric “Hall bars”, the theory is staged in geometries such as torus, sphere or plane, that avoid the complication of confining potentials. In a strong field, a boundary produces an edge current in a basically 1D layer whose properties differ from the bulk. Experiments on Hall bars with a close contact of edges, measured fluctuations of the tunnelling current (quantum shot noise) of fractional charge carriers [6].

Integer and fractional Hall effects are currently explained in two different ways. The former is caused by the presence of impurities, which remove the degeneracy and broaden Landau levels into bands of localised states centred at the Landau energies, which remain conducting. As electrons are added, no variation of the current is seen (the plateau), until the Fermi energy crosses the next conducting Landau energy, and the current has a jump.

Fractional plateaux are explained as an effect of the Coulomb interaction, but a final theory has not yet been given. The Nobel prize winner Laughlin [13] proposed the following ground states (symmetric gauge) at fillings $\nu = 1/q$:

$$\Psi_{1/q}(z_1, \dots, z_N) = \exp\left(-\sum_{i=1}^N \frac{|z_i|^2}{2\ell^2}\right) \prod_{j<i} (z_i - z_j)^q \quad (16.2)$$

The state is not normalised, $z_i = x_i - iy_i$, ℓ is the magnetic length ($\ell^2 = eB/hc$), q is odd for fermions. It is an eigenstate of the angular momentum and it is zero when particles collide, thus keeping the Coulomb energy at bay. Although not exact, numerical diagonalization with small N gives an impressive overlap with the true ground state. Haldane, Halperin and MacDonald obtained a hierarchy of ground states by adding quasi-holes to Laughlin states $\Psi_{1/q}$, obtaining the filling fractions $\nu = 2p/(2pq \pm 1)$. The process can be repeated, and in principle any value of filling fraction can be attained.

Another widely accepted scenario is the formation of “composite fermions”, i.e. electrons dressed with $2p$ units of magnetic flux. The model was proposed by Jain, and extended by Halperin, Lee and Read [9]. In a field B , the effective magnetic field felt by composites with electron density n is $B^* = B - 2pn(hc/e)$. The effective filling fraction is $\nu^* = n(hc/eB^*)$. For integer ν^* the IQHE of composite fermions occurs, corresponding to filling fractions $\nu = \nu^*/(2p\nu^* \pm 1)$.

In the Landau gauge, the Hilbert space spanned by the eigenstates of the lowest LL, consists of functions $\exp(-\sum_{i=1}^N \frac{|z_i|^2}{2\ell^2})f(z_1, \dots, z_N)$. The exponential factor is included in the measure of a Bargmann space of analytic functions f . With this separation, a Laughlin state $\psi_{1/q}$ is the power of a Vandermonde determinant $\Delta(z) = \prod_{j>k} (z_j - z_k)$. This and several other functions, such as Read-Moore and Read-Rezayi, are exact eigenstates of the Laplace-Beltrami operator, for the excited states of the Calogero-Sutherland Hamiltonian (an exactly solvable 1D many-particle model). They share the property of factoring into a Vandermonde determinant and a symmetric Jack polynomial. This intriguing link with FQHE was uncovered in 2008 by Bernevig and Haldane [4].

Another route was attempted by Tao and Thouless in 1983 [19], to explain the phase diagram as an effect of a Wigner crystal arrangement of electrons. However this possibility was ruled out because it predicts long-range solid order which is not observed in experiments. At the same time, the success of Laughlin's ansatz (for instance in giving a justification for the odd-denominator rule) discouraged further investigations in this direction.

The approach was revived in 2008 by Bergholtz and Karlhede [3], who showed that a crystal structure may arise in the strongly anisotropic quasi-one dimensional limit known as “thin torus” limit. The result was obtained by mapping the microscopic model in the thin torus limit to a 1D lattice gas model, that was analytically solved by Hubbard [8] and by Bak and Bruinsma [1]. The lattice gas has a fractal phase diagram, with plateaux of the density as a function of the chemical potential, forming a “Devil’s staircase”. Its implications for the FQHE were investigated by Rotondo et al. [14] and are in qualitative accord with the experimental diagram. To our knowledge, it is the first attempt to produce a realistic phase diagram.

Here, we review the thin torus limit and the phase diagram discussed in [14]. We then set the stage for the next result, with an introduction to the Calogero-Sutherland model and the Laplace-Beltrami operator, whose eigenfunctions are Jack polynomials.

There is consensus that the ground states of FQHE are adiabatically connected to the thin torus eigenstates. This is made particularly evident in the work [7], where a unitary operator is obtained that maps Slater determinants (or permanents) of a simpler diagonal theory, to the fully interacting eigenstates, such as the Laughlin states and, more generally, Jack-like states. The unitary operator, in essence, is the Dyson T-exp of the non-diagonal two-particle interaction in the Laplace-Beltrami operator.

16.2 Thin Torus and the Devil’s Staircase

The Hamiltonian for the FQHE in the lowest Landau level is:

$$H = \frac{\hbar\omega_c}{2} \sum_s a_s^\dagger a_s + \frac{1}{2} \sum_{\mathbf{s}} V(\mathbf{s}) a_{s_1}^\dagger a_{s_2}^\dagger a_{s_4} a_{s_3} \quad (16.3)$$

the operators a_s^\dagger and a_s create and destroy an electron in eigenstates $s = 1 \dots g$ spanning the lowest Landau level. $V(\mathbf{s}) = \langle s_1, s_2 | v | s_3, s_4 \rangle$ is the Coulomb matrix element. A background of positive charges cancels the Hartree term $s_1 = s_3, s_2 = s_4$. Note that the kinetic energy per particle $\hbar\omega_c/2$ can be identified with a chemical potential, that depends on B .

The earliest numerical calculations, for various filling fractions, were done by Yoshioka et al. [23, 25]. They considered the geometry of a periodic array of rectangles (torus) with commensurate area: $L_x L_y = 2\pi\ell^2 g$, where g is a large natural number. With the g degenerate eigenstates of the lowest LL in a rectangle, one constructs a basis of quasi-periodic eigenstates (Jacobi theta functions):

$$\theta_s(x, y) = \sum_{m \in \mathbb{Z}} \frac{1}{\sqrt{\ell L_y} \sqrt{\pi}} \exp \left[-i \frac{2\pi}{L_y} (s + mg)y - \frac{1}{2\ell^2} \left(x + mL_x - \frac{2\pi}{L_y} s \ell^2 \right)^2 \right]$$

with $0 \leq s \leq g - 1$; the normalisation is such that $\int_R dx dy \overline{\theta_s(x, y)} \theta_{s'}(x, y) = \delta_{ss'}$.

An electron in the rectangle R interacts with the electrons in R as well as with their copies. The Coulomb interaction is a periodic function of the lattice:

$$v(\mathbf{r}) = \sum_{\mathbf{m} \in \mathbb{Z}^2} \frac{e^2}{\sqrt{(x + m_x L_x)^2 + (y + m_y L_y)^2}} = \frac{1}{L_x L_y} \sum_{\mathbf{q}} \frac{2\pi e^2}{|\mathbf{q}|} e^{i\mathbf{q} \cdot \mathbf{r}}$$

where $q_x = \frac{2\pi}{L_x} n_x$ and $q_y = \frac{2\pi}{L_y} n_y$. The Coulomb matrix elements $V(\mathbf{s})$ are:

$$\begin{aligned} V(\mathbf{s}) &= \int_{R^2} d\mathbf{r}_1 d\mathbf{r}_2 \overline{\theta_{s_1}(\mathbf{r}_1)} \overline{\theta_{s_2}(\mathbf{r}_2)} v(\mathbf{r}_1 - \mathbf{r}_2) \theta_{s_3}(\mathbf{r}_1) \theta_{s_4}(\mathbf{r}_2) \\ &= \frac{1}{L_x L_y} \sum_{\mathbf{q}} v(\mathbf{q}) I_{s_1, s_3}(\mathbf{q}) I_{s_2, s_4}(-\mathbf{q}) \end{aligned}$$

where $I_{s, s'}(\mathbf{q}) = \int_R d\mathbf{r} \overline{\theta_s(\mathbf{r})} \theta_{s'}(\mathbf{r}) \exp(i\mathbf{q} \cdot \mathbf{r})$. The exact formula (2.9) in Yoshioka's paper [23] is obtained:

$$V(\mathbf{s}) = \frac{\delta_{s_1+s_2, s_3+s_4}}{L_x L_y} \sum_{\mathbf{q}} \frac{2\pi e^2}{|\mathbf{q}|} e^{-\frac{\ell^2}{2} |\mathbf{q}|^2 + i q_x \frac{2\pi \ell^2}{L_y} (s_3 - s_2)} \delta'(s_1 - s_3 + n_y) \quad (16.4)$$

The periodic (with period g) delta function δ' is used to sum on q_y :

$$= 2\pi e^2 \frac{\delta_{s_1+s_2, s_3+s_4}}{L_x L_y} \sum_{q_x} e^{-\frac{\ell^2}{2} q_x^2 + i q_x \frac{2\pi \ell^2}{L_y} (s_3 - s_2)} \sum_{m=-\infty}^{\infty} \frac{e^{-\frac{\ell^2}{2} \frac{4\pi^2}{L_y^2} (s_3 - s_1 + mg)^2}}{\sqrt{q_x^2 + \frac{4\pi^2}{L_y^2} (s_3 - s_1 + mg)^2}}$$

Now, approximate $\sum_{q_x} \approx \frac{L_x}{2\pi} \int dq_x$ and neglect terms $m \neq 0$ because of the exp factor. Equation (3) in the Tao and Thouless paper [19], is obtained:

$$V(\mathbf{s}) = \delta_{s_1+s_2, s_3+s_4} \frac{e^2}{L_y} \int_{-\infty}^{\infty} dq e^{-\frac{\ell^2}{2} q^2 + i q \frac{2\pi \ell^2}{L_y} (s_3 - s_2)} \frac{e^{-\frac{2\pi^2 \ell^2}{L_y^2} (s_3 - s_1)^2}}{\sqrt{q^2 + \frac{4\pi^2}{L_y^2} (s_3 - s_1)^2}} \quad (16.5)$$

A great simplification occurs in the thin torus limit $L_x \ll \ell$, when dependence on $s_3 - s_2$ disappears, and the matrix element only depends on $|s_3 - s_1|$ (modulo g). In this limit the integral yields a Bessel function:

$$V(|s_1 - s_3|) = \frac{e^2}{L_y} e^{-\frac{\pi^2 \ell^2}{L_y^2} (s_1 - s_3)^2} K_0 \left(\frac{\pi^2 \ell^2}{L_y^2} (s_1 - s_3)^2 \right)$$

The Coulomb operator becomes $\frac{1}{2} \sum_{s, t, u} V(|s - u|) a_s^\dagger a_t^\dagger a_{s+t-u} a_u$. The Fourier transform $c_k = \frac{1}{\sqrt{g}} \sum_{s=1}^g \exp(-\frac{2\pi i}{g} k s) a_s$ then makes the Hamiltonian diagonal:

$$H = \frac{\hbar\omega}{2} \sum_k n_k + \frac{1}{2} \sum_{k,k'=1}^g \tilde{V}(k-k') n_k n_{k'} \quad (16.6)$$

where $n_k = c_k^\dagger c_k$ and $\tilde{V}(k) = \sum_{s=1}^g V(s) \exp(\frac{2\pi i}{g} ks)$. When $g \gg$ range of $\tilde{V}(k)$, the problem of the ground state was exactly solved by Hubbard, in the context of classical 1D lattice gases [8]. He was searching for the sequence $\{n_k\} = \{0, 1\}^L$ minimising the energy $\sum_{k,k'} \tilde{V}(k-k') n_k n_{k'}$ at fixed density $\rho = \sum_k n_k / L$. Remarkably, Hubbard found that the answer is independent of the form of the interaction, provided the potential is convex with a finite asymptotic value, namely $\tilde{V}(k) \searrow 0$, $\tilde{V}(k+1) + \tilde{V}(k-1) > 2\tilde{V}(k)$. If the irreducible-fraction representation of the density is $\rho = p/q$, the ground-state sequence is periodic of period q , and is such that the position of the n -th particle is $\lfloor nq/p \rfloor$, where $\lfloor \cdot \rfloor$ is the floor function. Hubbard's periodic solution gives the vector of periodic occupation numbers of Landau states θ_s . In coordinate space, it is a Slater determinant of Jacobi theta functions. Minimising the Hamiltonian (16.6) is equivalent to finding which of the Hubbard states is the ground state of the energy $E = \frac{1}{2} \sum_{k,k'} \tilde{V}(k-k') n_k n_{k'} - \mu \sum_k n_k$ as a function of the chemical potential μ . This was solved by Bak and Bruinsma [1], who showed that the ground-state density $\rho(\mu)$ has a very interesting behaviour: it is a complete Devil's staircase, i.e. the Cantor function with plateaux at every rational number. (Recently, a related fractal hierarchy was discovered in the out of equilibrium counterpart of the model studied by Rotondo et al. [15].)

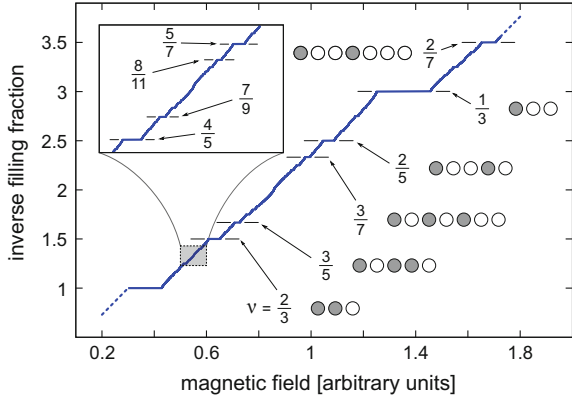
While the sequence of zeros and ones is independent of the potential, the widths of the plateaux do depend on it, and are given by the formula [1, 5]:

$$\Delta\mu = 2q \sum_{k=1}^{\infty} k [\tilde{V}(kq+1) + \tilde{V}(kq-1) - 2\tilde{V}(kq)].$$

To export the phase diagram of the lattice gas (μ, ρ) to the thin torus FQHE diagram $(B, 1/\nu)$ one has to rescale the Hamiltonian to cancel the dependence on B of the interaction term (of course, a global factor does not change the ground state). Only the rescaled kinetic energy depends on the magnetic field, and corresponds to μ in the lattice gas. The outcome is a diagram, Fig. 16.2, where the plateaux widths qualitatively reproduce the experimental observations. Notice that, while fractions with the same denominator have equal widths in μ , this symmetry is lost in the $(B, 1/\nu)$ plot. Another feature modifies the diagram: even-denominators thin torus ground states are forbidden because of Fermi statistics and modular invariance, as recently shown by Seidel [16].

Although the plot of the Devil's staircase in the thin torus limit is an intriguing result, we know that thin torus (TT) ground states can not describe truly two-dimensional FQH states, which are much better approximated by Laughlin's and composite fermion wavefunctions. However, this result encouraged us to investigate more in detail the algebraic structure of Laughlin' wavefunctions, that recently Haldane and Bernevig found to belong to a class of special orthogonal polynomials, the Jack polynomials. In the following we introduce this fascinating mathematical topic,

Fig. 16.2 The Devil's staircase for the FQHE in the thin torus limit. It shows a correct trend in the asymmetries of widths for particle-antiparticle exchange



introducing these polynomials as exact eigenfunctions of a well known integrable quantum system: the Calogero-Sutherland model.

16.3 Calogero-Sutherland and Laplace-Beltrami Operators

Consider N particles on the unit circle, that may cross and interact via a two-particle potential $1/r^2$, where r is their chordal distance [18]:

$$H_{CS}(\beta) = - \sum_{k=1}^N \frac{\partial^2}{\partial \theta_k^2} + \frac{1}{2} \beta(\beta - 1) \sum_{j < k} \frac{1}{\sin^2 \frac{1}{2}(\theta_j - \theta_k)} \quad (16.7)$$

$\beta > 0$ is a free parameter. The total momentum $P = -i \sum_k \partial / \partial \theta_k$ is conserved. Since $H_{CS} = \sum_{k=1}^N A_k^\dagger A_k + E_0$ where

$$A_k = -i \frac{\partial}{\partial \theta_k} + i \frac{\beta}{2} \sum_{j \neq k} \cotg \frac{1}{2}(\theta_k - \theta_j), \quad E_0 = \frac{\beta^2}{12} N(N^2 - 1), \quad (16.8)$$

the ground state is obtained by solving $A_k \psi_0 = 0$ for all $k = 1 \dots N$, and has energy E_0 . The extension beyond the sector $\theta_1 < \dots < \theta_N$ brings a problem of sign. For bosons or fermions:

$$\psi_0(\theta_1, \dots, \theta_N) = \prod_{j > k} |\sin \frac{1}{2}(\theta_j - \theta_k)|^\beta \times \begin{cases} 1 \\ \text{sign}(\theta_j - \theta_k) \end{cases}$$

In the variables $z_k = \exp(i\theta_k)$ the Hamiltonian and the total momentum become

$$H_{CS} = \sum_{k=1}^N (z_k \partial_k)^2 - 2\beta(\beta - 1) \sum_{j < k} \frac{z_j z_k}{(z_j - z_k)^2}, \quad P = \sum_{k=1}^N z_k \partial_k \quad (16.9)$$

Note that $\prod_{j>k} \sin^2 \frac{1}{2}(\theta_j - \theta_k) = \frac{\Delta(z)\Delta(1/z)}{2^{N(N-1)}}$ where $\Delta(z)$ is the Vandermonde determinant. Since $\bar{z} = 1/z$, the expression coincides with $|\Delta(z)|^2$ up to a constant. The bosonic ground state is

$$\psi_0(z) = |\Delta(z)|^\beta \tag{16.10}$$

with null total momentum. The excited states are searched in the factored form $|\Delta(z)|^\beta f(z)$. The eigenvalue equation $\psi_0^{-1} H_{CS} \psi_0 f = E f$ becomes $H_{LB} f = (E - E_0) f$, where the Laplace Beltrami operator is:

$$H_{LB} = \sum_{k=1}^N (z_k \partial_k)^2 + \frac{\beta}{2} \sum_{j \neq k} \frac{z_k + z_j}{z_k - z_j} (z_k \partial_k - z_j \partial_j) \tag{16.11}$$

The total momentum is $P' f = \psi_0^{-1} P \psi_0 f = \sum_k z_k \partial_k f = P f$ and commutes with H_{LB} . When promoting z from the unit circle to the whole complex plane, some differences arise. It is instructive to summarize the problem.

In the Bargmann space of entire functions $f(z)$, $z = (z_1 \dots z_N)$, such that $\int \frac{d^{2N}z}{\pi^N} e^{-\sum_k |z_k|^2} |f(z)|^2 < \infty$, we look for a potential V and energy value E such that $[\sum_{k=1}^N (z_k \partial_k)^2 + V(z_1 \dots z_N) - E] \Delta(z)^\beta = 0$, i.e. $E(\beta) - V(\mathbf{z}) = \frac{1}{\Delta^\beta} \sum_{k=1}^N (z_k \partial_k)^2 \Delta^\beta$. We recover the Hamiltonian H_{CS} of the Calogero-Sutherland model, (16.9), with eigenvalue $E(\beta) = \frac{1}{2} \beta^2 N(N-1)(2N-1)$.

If the excited states are searched with the factored form $\psi(\mathbf{z}) = \Delta(\mathbf{z})^\beta f(\mathbf{z})$, then f is an eigenfunction of the operator

$$\begin{aligned} [H' - E(\beta)]f &= \left[\Delta(z)^{-\beta} \sum_{k=1}^N (z_k \partial_k)^2 \Delta(z)^\beta + V(z) - E(\beta) \right] f(\mathbf{z}) \\ &= H_{LB} f(\mathbf{z}) + \frac{1}{2} \beta(N-1) P f(\mathbf{z}) \end{aligned} \tag{16.12}$$

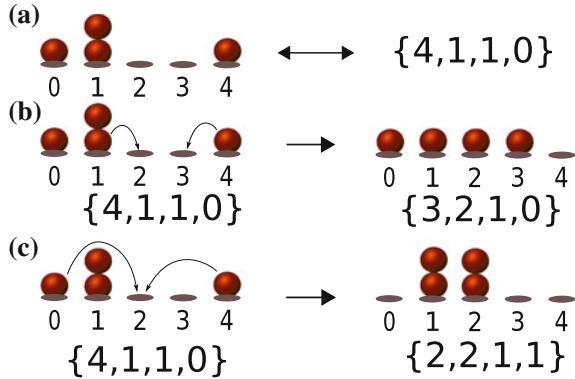
We obtained the Laplace-Beltrami operator (16.11), corrected by a boost. In Bargmann space, the operators H' and H_{LB} are not Hermitian, but they gain a nice mathematical property: in the basis of monomials of N variables, they are lower triangular. This allows to find all the eigenvectors, the celebrated Jack polynomials, that we introduce in the next section.

16.4 Jack Polynomials

The theory of Jack polynomials requires the concepts of partition and squeezing of a partition. They will fit naturally in the Fock space formalism.

Partitions are sequences $\lambda = \{\lambda_1, \dots, \lambda_N\}$ of integers $\lambda_1 \geq \dots \geq \lambda_N \geq 0$. The length $|\lambda|$ of a partition is the sum of its integers. A partition can be specified by the sequence of multiplicities of the integers. Physicists think of them as occupation

Fig. 16.3 **a** A configuration of occupancies corresponds to a partition. Here the lattice configuration on the left corresponds to the partition $\{4, 1, 1, 0\}$. **b** and **c** Lattice and partition representation of two of the possible squeezings of the partition $\{4, 1, 1, 0\}$



numbers (n_0, n_1, \dots) of particles on a semi-infinite lattice. For instance, $\{4, 1, 1, 0\}$ has 1 zero ($n_0 = 1$), 2 ones ($n_1 = 2$), and 1 four ($n_4 = 1$); then it can be written as $(120010 \dots) \equiv (12001)$ i.e. four particles in positions 0,1,1 and 4 (see Fig. 16.3a).

We define the squeezing of a partition in the picture of occupation numbers. A squeezing produces a new partition where two particles are moved toward each other. A squeezing of (12001) is (11110) (one of the two particles in site 1 and the particle in site 4 are moved to sites 2 and 3). There are two other squeezes of (12001) : (03010) (we move the particle in site 0 and the one in site 4 by one or three sites) and (02200) (we move 0 and 4 by two sites). A pictorial representation of some of these squeezings is given in Fig. 16.3b, c.

We define a partial order among partitions with same number of particles: $\lambda < \mu$ if the partition μ can be obtained by one or more squeezings of the partition λ .

A partition $\lambda = \{\lambda_1, \dots, \lambda_N\}$ specifies a monomial $\mathbf{m}_\lambda(\mathbf{z})$, $\mathbf{z} \in \mathbb{C}^N$:

$$\mathbf{m}_\lambda(\mathbf{z}) = \sum_{\pi} z_{\pi_1}^{\lambda_1} \dots z_{\pi_N}^{\lambda_N} \tag{16.13}$$

The sum covers permutations π of $1, \dots, N$ without repetition of equal terms. A monomial is a symmetric polynomial of degree $|\lambda|$. For example: $\mathbf{m}_{3,1,1}(\mathbf{z}) = z_1^3 z_2 z_3 + z_2^3 z_1 z_3 + z_3^3 z_1 z_2$. The symmetric monomials $\mathbf{m}_\lambda(\mathbf{z})$, $\mathbf{z} \in \mathbb{C}^N$, are an orthogonal basis in Bargmann space.

Theorem 1 (Sogo [17]) *In the basis of monomials with fixed degree, the Laplace-Beltrami operator H_{LB} is triangular, i.e.:*

$$H_{LB} \mathbf{m}_\lambda(\mathbf{z}) = E_\lambda \mathbf{m}_\lambda(\mathbf{z}) + \sum_{\mu < \lambda} C_{\lambda\mu} \mathbf{m}_\mu(\mathbf{z})$$

with diagonal values (the eigenvalues of H_{LB}):

$$E_\lambda = \sum_{j=1}^N \lambda_j^2 + \beta \sum_{j < k} (\lambda_j - \lambda_k) = \sum_{j=1}^N \lambda_j^2 + \beta(N + 1 - 2j)\lambda_j \tag{16.14}$$

$C_{\lambda\mu} = 0$ if μ is not a partition squeezed from λ ,

$$C_{\lambda\mu} = 2\beta(\lambda_i - \lambda_j) \binom{n(\mu_i)}{2} \quad (\mu_i = \mu_j), \quad C_{\lambda\mu} = 2\beta(\lambda_i - \lambda_j)n(\mu_i)n(\mu_j) \quad (\mu_i \neq \mu_j)$$

where the squeezing takes $(\lambda_i, \lambda_j) \rightarrow (\mu_i, \mu_j)$ ($i < j$).

The diagonalization of the matrix brings the polynomial eigenfunctions of H_{LB} of degree $|\lambda|$: the symmetric Jack polynomials $J_\lambda(\mathbf{z})$.

$$J_\lambda(\mathbf{z}) = m_\lambda(\mathbf{z}) + \sum_{\mu < \lambda} b_{\lambda,\mu} m_\mu(\mathbf{z}) \tag{16.15}$$

Lapointe and Lascaux [12] gave a determinantal expression for J_λ (with $b_{\lambda\lambda} \neq 1$).

Example $N = 3, |\lambda| = 3$.

The partitions with $N = 3$ particles and length 3 are $\{1, 1, 1\}, \{2, 1, 0\}, \{3, 0, 0\}$. Then, the matrix H_{LB} is 3×3 . The basis of monomials are: $\mathbf{m}_{111} = z_1 z_2 z_3, \mathbf{m}_{210} = z_1^2 z_2 + z_1^2 z_3 + z_2^2 z_1 + z_2^2 z_3 + z_3^2 z_1 + z_3^2 z_2, \mathbf{m}_{300} = z_1^3 + z_2^3 + z_3^3$.

$$H_{LB} \begin{bmatrix} m_{111} \\ m_{210} \\ m_{300} \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 12\beta & 5 + 4\beta & 0 \\ 0 & 6\beta & 9 + 6\beta \end{bmatrix} \begin{bmatrix} m_{111} \\ m_{210} \\ m_{300} \end{bmatrix}$$

Diagonalization brings out 3 Jack polynomials: $J_{111} = m_{111}, J_{210} = m_{210} + \frac{6\beta}{1+2\beta} m_{111}$, and $J_{300} = m_{300} + \frac{3\beta}{2+\beta} m_{210} + \frac{6\beta^2}{(1+\beta)(2+\beta)} m_{111}$ with eigenvalues $3, 5 + 4\beta, 9 + 6\beta$.

16.5 Fock Space

Let us introduce the Fock space formalism for the lowest Landau level. If $|\lambda\rangle, \lambda = 0, 1, 2, \dots$ is an orthogonal basis for the single-particle Hilbert space, the normalized basis is $\frac{1}{\sqrt{\nu_\lambda}} |\lambda\rangle$, with $\nu_\lambda = \langle \lambda | \lambda \rangle$.

For the lowest Landau level: $\langle z | \lambda \rangle = z^\lambda, \nu_\lambda = \lambda!, \lambda = 0, 1, \dots$

The orthogonal bases for the Hilbert space of N bosons or fermions are given respectively by:

$$\begin{aligned} |\mathbf{m}_\lambda\rangle &= \sum_{\pi} |\lambda_{\pi_1} \dots \lambda_{\pi_N}\rangle, \\ |\mathbf{s}_\lambda\rangle &= \sum_{\pi} (-)^{\pi} |\lambda_{\pi_1} \dots \lambda_{\pi_N}\rangle, \end{aligned} \tag{16.16}$$

where $\lambda = \{\lambda_1, \dots, \lambda_N\}$, with $\lambda_1 \geq \dots \geq \lambda_N$, is a partition which specifies the single particle states (for fermions equality is forbidden). The sum runs over permutations of indices $\{1, \dots, N\}$ without repetition of equal terms and $(-)^{\pi}$ is the parity of the permutation. For example:

$$\begin{aligned} |m_{440}\rangle &= |440\rangle + |404\rangle + |044\rangle, \\ |sl_{410}\rangle &= |410\rangle - |401\rangle + |041\rangle - |014\rangle + |104\rangle - |140\rangle. \end{aligned} \quad (16.17)$$

We recover the many-particle wavefunctions of lowest Landau level, in Bargmann space: $\langle \mathbf{z} | m_{\lambda} \rangle = m_{\lambda}(\mathbf{z})$ (the monomials in (16.13)) and $\langle \mathbf{z} | sl_{\lambda} \rangle = sl_{\lambda}(\mathbf{z})$ (Slater determinants).

The bases (16.16) have the equivalent notations:

- $|\lambda_1, \dots, \lambda_N\rangle$, specifying the single-particle quantum numbers ($\lambda_1 \geq \dots \geq \lambda_N$);
- $|n_0, \dots, n_{\infty}\rangle$, specifying how many particles share the same quantum number.

We introduce the canonical algebra of creation and destruction operators of 1-particle states $\frac{1}{\sqrt{\nu_{\lambda}}} |\lambda\rangle$, for bosons (+) and fermions (-):

$$\begin{aligned} a_{\lambda}^{\dagger} | \dots n_{\lambda} \dots \rangle &= (\pm 1)^{n_0 + \dots + n_{\lambda-1}} \sqrt{n_{\lambda} + 1} | \dots n_{\lambda} + 1 \dots \rangle, \\ a_{\lambda} | \dots n_{\lambda} \dots \rangle &= (\pm 1)^{n_0 + \dots + n_{\lambda-1}} \sqrt{n_{\lambda}} | \dots n_{\lambda} - 1 \dots \rangle, \\ a_{\lambda} | \dots n_{\lambda} \dots \rangle &= 0 \quad \text{if } n_{\lambda} = 0. \end{aligned} \quad (16.18)$$

The operators allow for an efficient description of the squeezing operation introduced before, which is implemented by the following operator:

$$s_{u,m,k} = \sqrt{\frac{\nu_{u-k}\nu_{m+k}}{\nu_u\nu_m}} a_{u+k}^{\dagger} a_{m-k}^{\dagger} a_m a_u \quad (16.19)$$

for $0 \leq u < m$ and $0 < k < m - u$. Its action on the basis $|m_{\lambda}\rangle$ is:

$$s_{u,m,k} |m_{\lambda}\rangle = \begin{cases} 0 & \text{if } u \notin \lambda \text{ or } m \notin \lambda \\ (n_{u+k} + 2)(n_{u+k} + 1) |m_{\mu}\rangle & k = \frac{m-u}{2} \in \mathbb{N} \\ (n_{u+k} + 1)(n_{m-k} + 1) |m_{\mu}\rangle & \text{otherwise,} \end{cases} \quad (16.20)$$

where μ is obtained from λ by substituting two particles of quantum numbers u and m with two particles of quantum numbers $u+k$ and $m-k$, hence by ‘‘squeezing two particles by k ’’. The action on $|sl_{\lambda}\rangle$ is

$$s_{u,m,k} |sl_{\lambda}\rangle = (-)^{N_{sw}} |sl_{\mu}\rangle, \quad (16.21)$$

where N_{sw} is the number of exchanges restoring the decreasing order of the sequence.

16.6 A Map for FQHE Effective States

In analogy with Sogo’s theorem, establishing a correspondence between monomials and Jack polynomials, we obtain an operator that generates symmetric or antisymmetric eigenstates of the Laplace-Beltrami operator, from simple root states. Some eigenstates are effective ground states of the FQHE.

Let H_0 be a Hermitian operator, expressed solely in terms of number operators. Its eigenstates are the vectors $|\mathbf{m}_\lambda\rangle$ (bosons) or $|\mathbf{s}_\lambda\rangle$ (fermions). We collectively denote them as $|\lambda\rangle$. Let $H = H_0 + V$, where V is a linear combination of squeezing operators. Note that if $|\lambda\rangle$ is an eigenstate of H_0 with a finite number of particles, then there is n such that V^n annihilates $|\lambda\rangle$.

Theorem 2 *If E_λ is a non degenerate eigenvalue of H_0 , with eigenvector $|\lambda\rangle$ then E_λ is an eigenvalue of H with eigenvector*

$$|\psi_\lambda\rangle = \sum_{k=0}^n [(E_\lambda - H_0)^{-1} V]^k |\lambda\rangle \tag{16.22}$$

Proof Multiplication by $(E_\lambda - H_0)^{-1} V$ gives the identity

$$(E_\lambda - H_0)^{-1} V |\psi_\lambda\rangle = |\psi_\lambda\rangle - |\lambda\rangle \tag{16.23}$$

Multiplication by $(E_\lambda - H_0)$ gives $V |\psi_\lambda\rangle = (E_\lambda - H_0) |\psi_\lambda\rangle$.

The reverse is true: if $|\psi\rangle$ is an eigenstate of H with non-degenerate eigenvalue E , then $|\lambda\rangle = |\psi\rangle - (E - H_0)^{-1} V |\psi\rangle$ is an eigenvector of H_0 with same eigenvalue.

By construction $|\psi_\lambda\rangle$ is a linear combination of basis states that are squeezed from $|\lambda\rangle$: $|\psi_\lambda\rangle = \sum_{\mu \leq \lambda} b_{\lambda\mu} |\mu\rangle$. Equation (16.23) gives $b_{\lambda\lambda} = 1$ and the recursive relation:

$$b_{\lambda\mu} = \frac{1}{E_\lambda - E_\mu} \sum_{\mu < \mu' < \lambda} \langle \mu | V | \mu' \rangle b_{\lambda, \mu'} \quad (\mu \neq \lambda) \tag{16.24}$$

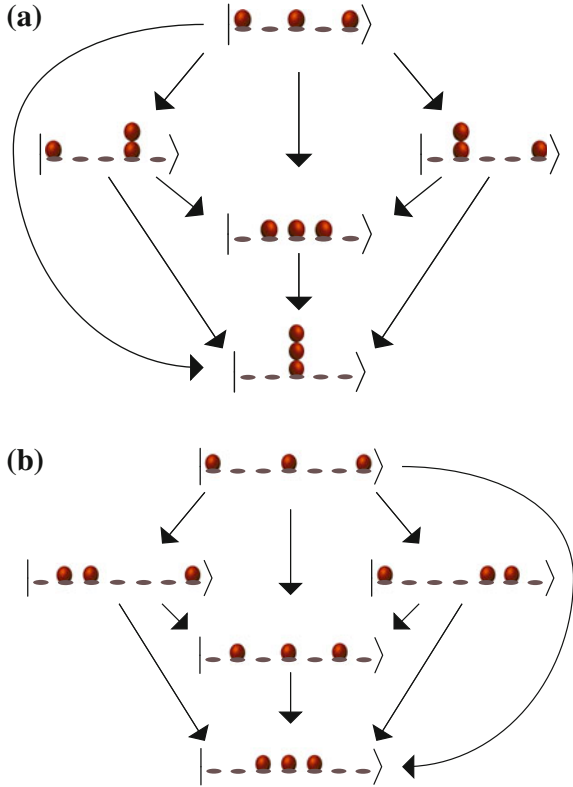
The sum only involves states that yield $|\mu\rangle$ after a single squeezing contained in V and that descend from the root $|\lambda\rangle$ by one or more squeezings.

We can use this theorem to obtain the FQHE effective states that are a Jack polynomial times a Vandermonde determinant, since such wavefunctions are eigenstates of the “Generalized Laplace Beltrami” operator:

$$H = \sum_{k=1}^N (z_k \partial_k)^2 + K \sum_{j \neq k} \frac{z_k + z_j}{z_k - z_j} (z_k \partial_k - z_j \partial_j) - K \sum_{j \neq k} \frac{z_j^2 + z_k^2}{(z_j - z_k)^2} (1 - \pi_{jk}) \tag{16.25}$$

where π_{jk} is the exchange operator of particles j, k . In the bosonic sector $\pi_{jk} = 1$, and we recover the Laplace Beltrami operator (16.11). $K = (1 - 2q \pm 1)/(p + 1)$ (+1 for bosons, -1 for fermions), with filling fraction $\nu = p/q$.

Fig. 16.4 The repeated action of the squeezing operator in (16.26) on a bosonic (a) and a fermionic (b) state with three particles. Each state is specified by the occupancy numbers, and each arrow represents the action of a squeeze. The operator in (16.26) automatically takes into account the possibility of multiple occupancies of the bosonic case



By evaluating H in second quantization we obtained $H = D + KS$, where S is a linear combination of squeezing operators

$$S = \sum_{s=0}^{\infty} \sum_{t=1}^{\infty} \sum_{u=1}^{\infty} u \sqrt{\frac{(s+t)!(s+u)!}{(s+t+u)!s!}} a_{s+t}^{\dagger} a_{s+u}^{\dagger} a_{s+t+u} a_s \quad (16.26)$$

and D incorporates the kinetic term and the diagonal part of the two-body potential (up to an irrelevant additive constant for fermions):

$$D = \sum_{m=0}^{\infty} m^2 a_m^{\dagger} a_m + \frac{K}{2} \sum_{m' < m} (m - m') a_m^{\dagger} a_m a_{m'}^{\dagger} a_{m'}. \quad (16.27)$$

Remarkably, we obtain Laughlin's wavefunction at filling fraction $1/q$ if we use as root the thin torus ground state with same filling fraction. In Fig. 16.4 we show the action of the squeezing operator on the thin torus ground state for $N = 3$ and $\nu = 1/3$, generating Laughlin's wavefunction $q = 3$.

16.7 Conclusions

First, we showed that the FQHE Hamiltonian can be mapped, in the thin torus limit, to a one-dimensional (classical) lattice gas model, whose main ingredient is a long-range repulsive interaction. The map brings one to interpret the ground states as Hubbard states and to prove their incompressibility, i.e., the fact that they are stable in finite intervals of magnetic field. Importantly, this enables to produce a Devil's staircase phase diagram for the filling fractions, which shows realistic features, such as the asymmetry of the plateau widths, qualitatively in accordance with the experimental landscape.

Second, we obtained an explicit Fock-space representation of FQHE wavefunctions. The generic recursive relation (16.24) recovered the known relations for the coefficients of the expansion of Jack polynomials [20].

To address the question *what are Jacks doing on the Devil's staircase?*, it would be useful to identify in the open plane and in the symmetric gauge, a mechanism (analogous to the thin torus mapping) that selects the monomials or Slater determinants that are mapped by the operator (16.22), specified by (16.26) and (16.27), to the Laughlin state or others relevant for FQHE. Another intriguing question is to understand whether the composite fermion wavefunctions that generalise the Laughlin's ansatz at more complicated filling fractions, are related in some way to the theory of Jacks. This is already the case for Moore-Read and Read-Rezayi wavefunctions, and would be one more key indication about the fundamental role that Jack polynomials play in the theory of FQHE.

References

1. P. Bak, R. Bruinsma, One-dimensional Ising model and the complete Devil's staircase. Phys. Rev. Lett. **49**(4) (1982). <https://doi.org/10.1103/PhysRevLett.49.249>
2. W. Baratta, P.J. Forrester, Jack polynomial fractional quantum Hall states and their generalizations. Nucl. Phys. B **843**[PM], 362–381 (2011). <https://doi.org/10.1016/j.nuclphysb.2010.09.018>
3. E.J. Bergholtz, A. Karlhede, Quantum Hall systems in Tao-Thouless limit. Phys. Rev. B **77**, 155308 (2008). <https://doi.org/10.1103/PhysRevB.77.155308>
4. B.A. Bernevig, F. Haldane, Model fractional quantum Hall states and Jack polynomials. Phys. Rev. Lett. **100**, 246802 (2008). <https://doi.org/10.1103/PhysRevLett.100.246802>
5. S.E. Burkov, Y.G. Sinai, Phase diagrams of one-dimensional lattice models with long-range antiferromagnetic interaction. Russ. Math. Surv. **38**(4), 235 (1983). <https://doi.org/10.1070/RM1983v038n04ABEH004211>
6. R. de-Picciotto, M. Reznikov, M. Heiblum, V. Umansky, G. Bunin, D. Mahalu, Direct observation of a fractional charge. Nature **389**(6647), 162 (1997). <https://doi.org/10.1038/38241>
7. A. Di Gioacchino, L.G. Molinari, V. Erba, P. Rotondo, Unified Fock space representation of fractional quantum Hall states. Phys. Rev. B **95**, 245123 (8 pp.) (2017). <https://doi.org/10.1103/PhysRevB.95.245123>
8. J. Hubbard, Generalized Wigner lattices in one dimension and some applications to tetracyanoquinodimethane (TCNQ) salts. Phys. Rev. B **17**(2), 494–505 (1978). <https://doi.org/10.1103/PhysRevB.17.494>

9. J.K. Jain, *Composite Fermions* (Cambridge University Press, 2007)
10. K. v. Klitzing, G. Dorda, M. Pepper, New method for high-accuracy determination of the fine-structure constant based on quantized Hall resistance. *Phys. Rev. Lett.* **45**(6), 494–497 (1980). <https://doi.org/10.1103/PhysRevLett.45.494>
11. L. Lapointe, L. Vinet, Exact operator solution of the Calogero-Sutherland model. *Commun. Math. Phys.* **178**, 425–452 (1996). <https://doi.org/10.1007/BF02099456>
12. L. Lapointe, A. Lascaux, J. Morse, Determinantal expression and recursion for Jack polynomials. *Electron. J. Comb.* **7**(1) (7 pp.) (2000), <http://www.combinatorics.org/ojs/index.php/eljc/article/view/v7i1n1/pdf>
13. R.B. Laughlin, Anomalous quantum Hall effect: an incompressible quantum fluid with fractionally charged excitations. *Phys. Rev. Lett.* **50**, 1395 (1983). <https://doi.org/10.1103/PhysRevLett.50.1395>
14. P. Rotondo, L.G. Molinari, P. Ratti, M. Gherardi, Devil's staircase phase diagram of the fractional quantum Hall effect in the thin-torus limit. *Phys. Rev. Lett.* **116**, 256803 (5 pp.) (2016). <https://doi.org/10.1103/PhysRevLett.116.256803>
15. P. Rotondo, A.L. Sellerio, P. Glorioso, S. Caracciolo, M. Cosentino Lagomarsino, M. Gherardi, Current quantization and fractal hierarchy in a driven repulsive lattice gas, [arXiv:1708.02403](https://arxiv.org/abs/1708.02403) [cond-mat.stat-mech] (2017)
16. A. Seidel, S-duality constraints on 1D patterns associated with fractional quantum Hall states. *Phys. Rev. Lett.* **105**, 026802 (2010). <https://doi.org/10.1103/PhysRevLett.105.026802>
17. K. Sogo, Eigenstates of Calogero-Sutherland-Moser model and generalized Schur functions. *J. Math. Phys.* **35**(5), 2282–2296 (1994). <https://doi.org/10.1063/1.530552>
18. B. Sutherland, *Beautiful Models* (World Scientific, 2004)
19. R. Tao, D.J. Thouless, Fractional quantization of Hall conductance. *Phys. Rev. B* **28**(2), 1142–1144 (1983). <https://doi.org/10.1103/PhysRevB.28.1142>
20. R. Thomale, B. Estienne, N. Regnault, B.A. Bernevig, Decomposition of fractional quantum Hall model states: product rule symmetries and approximations. *Phys. Rev. B* **84**, 045127 (2011). <https://doi.org/10.1103/PhysRevB.84.045127>
21. D. C. Tsui, H. L. Stormer, A. C. Gossard, Two-dimensional magnetotransport in the extreme quantum limit. *Phys. Rev. Lett.* **48**, 1559 (1982). <https://doi.org/10.1103/PhysRevLett.48.1559>
22. R. Willett, J.P. Eisenstein, H.L. Stormer, D.C. Tsui, C. Gossard, J.H. English, Observation of an even-denominator quantum number in the fractional quantum Hall effect. *Phys. Rev. Lett.* **59**(15), 1776–1779 (1987). <https://doi.org/10.1103/PhysRevLett.59.1776>
23. D. Yoshioka, Ground state of the two-dimensional charged particles in a strong magnetic field and the fractional quantum Hall effect. *Phys. Rev. B* **29**(12), 6833–6839 (1984). <https://doi.org/10.1103/PhysRevB.29.6833>
24. D. Yoshioka, *The Quantum Hall Effect*, 2nd ed. Springer Series in Solid State Science, vol. 133 (Springer, 2002)
25. D. Yoshioka, B.I. Halperin, P.A. Lee, Ground state of two-dimensional electrons in strong magnetic fields and $\frac{1}{3}$ quantized Hall effect. *Phys. Rev. Lett.* **50**(16), 1219–1222 (1983). <https://doi.org/10.1103/PhysRevLett.50.1219>

Chapter 17

A Study of the Phase Diagram of Symmetric Binary Gaussian Mixtures



Davide Pini

Abstract We perform an analytic study of the main features of the phase diagram of symmetric binary Gaussian mixtures such that the range and strength of the interactions between particles of one species are the same as those of the other species. We focus on the relative locations of the spinodal and λ -lines, i.e., the boundaries beyond which the uniform mixture becomes unstable respectively towards bulk demixing and microphase formation. We find that, when the λ -line is present, three situations may occur: (i) The spinodal line does not exist, and the λ -line spans the whole concentration axis. (ii) Both the λ - and spinodal lines are present and span the whole concentration axis, but the spinodal instability is always preempted by the λ -line. (iii) The spinodal instability is the only one present at intermediate concentrations, but is preempted by the λ -line at high and low concentrations.

17.1 Introduction

In the coarse-grained description of polymeric solutions, polymers are replaced by effective particles with no internal degrees of freedom such that their mutual interactions depend solely on the distance r between the polymer centers of mass. In particular, it has long been acknowledged that repulsive, bounded pair potentials of Gaussian form provide a simple modelization of the interactions between polymer chains in an athermal solvent [1–3]. On the same footing, binary polymer mixtures can be represented as an assembly of particles of two species interacting by a two-body, repulsive, athermal Gaussian potential [2] given by

$$\beta v_{ij}(r) = \epsilon_{ij} e^{-(r/R_{ij})^2}, \quad (17.1)$$

D. Pini (✉)

Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano,
Via Celoria 16, 20133 Milan, Italy
e-mail: davide.pini@fisica.unimi.it

© Springer Nature Switzerland AG 2018

P. F. Bortignon et al. (eds.), *Toward a Science Campus in Milan*,
https://doi.org/10.1007/978-3-030-01629-6_17

209

where $i, j = 1, 2$ identify the particle species, $\epsilon_{ij} > 0$ and $R_{ij} > 0$ determine the interaction strengths and ranges respectively, and $\beta = 1/k_B T$, T being the absolute temperature and k_B the Boltzmann constant.

For suitable interaction parameters, binary Gaussian mixtures have been shown to display a demixing transition [4–8], whereby at high density the mixed phase separates into two mixtures with different concentrations. Within a simple mean-field approximation, demixing occurs provided ϵ_{ij} and R_{ij} satisfy the condition [4, 5]

$$\left(\frac{R_{11} R_{22}}{R_{12}^2} \right)^3 < \frac{\epsilon_{12}^2}{\epsilon_{11} \epsilon_{22}}. \quad (17.2)$$

Interestingly, this is not the only possible scenario: if particles of different species have a strong affinity for each other, not only is demixing inhibited but the mixture can even form ordered structures in order to take maximum advantage from such an affinity. This situation does not take place for the parameter choices appropriate to model linear polymer chains [2], but may occur e.g. in mixtures of dendrimers [9]. Another case leading to this scenario was considered in [10], and its detailed investigation [11] brought forth a very rich phase diagram featuring a number of microphases, i.e., periodic structures with density modulations much larger than the size of the particles.

At the same time, the quantitative conditions on the interaction parameters for which these periodic phases can be expected have not, to our knowledge, been determined yet. In fact, for generic ϵ_{ij} , R_{ij} this problem appears rather daunting, even within a simple theory such as that used to obtain (17.2) for the demixing transition. However, it considerably simplifies if one contents oneself with symmetric mixtures, i.e., those for which $\epsilon_{11} = \epsilon_{22}$, $R_{11} = R_{22}$. Since, say, ϵ_{11} and R_{11} can be absorbed into the definition of the energy and length scales, this drastically reduces the free parameters to $\epsilon_{12}/\epsilon_{11}$ and R_{12}/R_{11} . Then, the criteria for the occurrence of periodic phases can be easily determined analytically. This is the purpose of the present paper.

Admittedly, the symmetric model represents an idealized situation which may not lend itself to describing actual polymer mixtures. However, we expect that the picture which will be brought forth in this study will occur also in the more general situation.

The paper is organized as follows: in Sect. 17.2 the conditions for the occurrence of periodic phases are determined and the shape of the stability boundary of the homogeneous phase is discussed; in Sect. 17.3 our results are summarized, and the stability boundaries are plotted for some specific cases in order to illustrate the scenarios established in the former section.

17.2 Existence and Behavior of the λ - and Spinodal Lines

We shall be studying a binary mixture of particles interacting by the athermal Gaussian potential $v_{ij}(r)$ given by (17.1) in the symmetric case such that $\epsilon_{11} = \epsilon_{22}$, $R_{11} = R_{22}$.

Since the mixture has been assumed to be athermal, its thermodynamic state is determined by the average densities of the two species $\rho_i \equiv N_i/V$, where N_i is the number of particles of species i and V is the volume of the system. If we denote by $\rho_i(\mathbf{r})$ the *local* density of species i at position \mathbf{r} , in a homogeneous state one has $\rho_i(\mathbf{r}) \equiv \rho_i$ identically. In order for this to be the actual equilibrium state, i.e., to correspond to the minimum of the Helmholtz free energy A at given N_i and V , the condition of stability with respect to small perturbations of the local density $\delta\rho_i(\mathbf{r})$ has to be satisfied. This amounts to the constraint

$$\sum_{i,j=1}^2 \int d^3\mathbf{r} \int d^3\mathbf{r}' c_{ij}(\mathbf{r}, \mathbf{r}') \delta\rho_i(\mathbf{r}) \delta\rho_j(\mathbf{r}') < 0, \quad (17.3)$$

where we have introduced the direct correlation function $c_{ij}(\mathbf{r}, \mathbf{r}')$. The latter is defined in terms of the free energy A regarded as a functional of the local densities via the expression

$$c_{ij}(\mathbf{r}, \mathbf{r}') \equiv \left. \frac{\delta^2(-\beta A)}{\delta\rho_i(\mathbf{r}) \delta\rho_j(\mathbf{r}')} \right|_{\rho_i, \rho_j}. \quad (17.4)$$

We remark that, according to (17.4), $c_{ij}(\mathbf{r}, \mathbf{r}')$ includes the ideal-gas contribution, at variance with the more widely adopted definition. For a homogeneous state, one has $c_{ij}(\mathbf{r}, \mathbf{r}') = c_{ij}(\mathbf{r} - \mathbf{r}')$, and (17.3) is equivalent to requiring that the matrix with elements $\tilde{c}_{ij}(\mathbf{k})$ be negative definite for every \mathbf{k} , where we have used the tilde to denote the Fourier transform with respect to \mathbf{r} :

$$\tilde{c}_{ij}(\mathbf{k}) = \int d^3\mathbf{r} e^{-i\mathbf{k}\cdot\mathbf{r}} c_{ij}(\mathbf{r}). \quad (17.5)$$

In the following, correlations will be described by the well-known random phase approximation (RPA) [12]. According to the RPA, $\tilde{c}_{ij}(k)$ is given by

$$\tilde{c}_{ij}(k) = -\frac{\delta_{ij}}{\rho_i} - \beta \tilde{v}_{ij}(k), \quad (17.6)$$

where k is the modulus of \mathbf{k} , and δ_{ij} is the Kroenecker delta. For the Gaussian potential (17.1), one has

$$\beta \tilde{v}_{ij}(k) = \pi^{3/2} R_{ij}^3 \epsilon_{ij} e^{-(R_{ij}k)^2/4}, \quad (17.7)$$

so that $\tilde{c}_{ij}(k) < 0$ for every k , and the above condition reduces to

$$\det ||\tilde{c}_{ij}(k)|| > 0 \quad \text{for every } k, \quad (17.8)$$

where \det indicates the determinant. A violation of this condition for a certain homogeneous state ρ_1, ρ_2 and wave vector k means that such a state is unstable with respect to density fluctuations with characteristic length $d \sim 2\pi/k$. The situation more commonly encountered corresponds to the spinodal instability found for $k=0$. In this case, the boundary between the thermodynamic states for which condition (17.8) is satisfied and those for which it is violated is the spinodal line, which is given by the locus of densities ρ_1, ρ_2 such that

$$\det ||\tilde{c}_{ij}(k=0)|| = 0. \quad (17.9)$$

From (17.4) it is readily found that the spinodal is the line in the ρ_1 - ρ_2 plane where the Hessian determinant of the Helmholtz free energy A vanishes. Inside the domain bounded by the spinodal, the Hessian of A is negative, and the homogeneous state is mechanically unstable with respect a perturbation of infinite wavelength, which corresponds to phase separation into two bulk phases with different ρ_i .

Here, on the other hand, we are mostly interested in the situation in which the parameters ϵ_{ij}, R_{ij} are chosen in such a way that the instability occurs at non-vanishing k . The boundary of the homogeneous phase is then the so-called λ -line. Requiring that the determinant of the matrix $||\tilde{c}_{ij}(k)||$ be vanishing is not sufficient to identify the λ -line, as it gives just one constraint on the three unknowns ρ_1, ρ_2 , and k . The additional requirement is that k must be chosen in such a way that (17.8) is indeed satisfied for *all* states lying on one side of the line. This leads to the conditions

$$\det ||\tilde{c}_{ij}(k)|| = 0, \quad (17.10)$$

$$\frac{d}{dk} \det ||\tilde{c}_{ij}(k)|| = 0 \quad (17.11)$$

for some $k \neq 0$.

In the following, the thermodynamic state will be expressed in terms of the total density $\rho = \rho_1 + \rho_2$ and concentration $\phi = \rho_1/\rho$. We also introduce a rescaled density y and concentration z , a rescaled wave vector q , and two dimensionless parameters H, α defined by

$$y = \pi^{3/2} \epsilon_{12} R_{12}^3 \rho \quad (17.12)$$

$$z = 2\phi - 1 \quad (17.13)$$

$$q = \frac{R_{12}}{\sqrt{2}} k \quad (17.14)$$

$$\alpha = \frac{R_{11}^2}{2R_{12}^2} \quad (17.15)$$

$$H = \frac{\epsilon_{11}R_{11}^3}{\epsilon_{12}R_{12}^3} \quad (17.16)$$

with $y > 0$, $-1 \leq z \leq 1$, $q > 0$, $\alpha > 0$, $H > 0$. We remark that, for the symmetric mixture in hand, the phase diagram is symmetric with respect to the equimolar line $z = 0$, so that both the spinodal and λ -lines must be even function of z .

By using (17.6) in (17.9) we obtain the equation for the spinodal line

$$y^2(1 - z^2)(H^2 - 1) + 4yH + 4 = 0. \quad (17.17)$$

This is solved immediately to give

$$y = 2 \frac{H + \sqrt{1 - (1 - H^2)z^2}}{(1 - H^2)(1 - z^2)} \quad H < 1, \quad (17.18)$$

where the condition on H stems from requiring that y must be real and positive. Hence, $H < 1$ is a necessary and sufficient condition for the existence of the spinodal. This condition is obtained here as a special case of (17.2). According to (17.18), the spinodal line on the y - z plane is a convex function of z which assumes its minimum $y_{\min} = 2/(1 - H)$ at $z = 0$, and diverges for $z \rightarrow \pm 1$. Again, (17.18) is a special case of the expression for generic, non-symmetric Gaussian mixtures established in [4, 5].

We now turn to the λ -line. Using again (17.6) in (17.10), (17.11) we find that along the λ -line one must have

$$y^2(1 - z^2)(H^2 e^{-2\alpha q^2} - e^{-q^2}) + 4yH e^{-\alpha q^2} + 4 = 0 \quad (17.19)$$

$$y^2(1 - z^2)(2\alpha H^2 e^{-2\alpha q^2} - e^{-q^2}) + 4\alpha yH e^{-\alpha q^2} = 0. \quad (17.20)$$

If we introduce the variable

$$x = H^2 e^{-(2\alpha-1)q^2} \quad (17.21)$$

we obtain the following expressions for y and z :

$$y = \frac{1 - 2\alpha x}{\alpha x + \alpha - 1} \left(\frac{H}{x^\alpha} \right)^{1/(2\alpha-1)} \quad (17.22)$$

$$z = \pm \frac{\sqrt{1 - 4\alpha^2 x}}{1 - 2\alpha x}, \quad (17.23)$$

which give the λ -line in the y - z plane parametrically as a function of x , or, equivalently, q . By imposing the conditions $y > 0$, $-1 < z < 1$ which identify the physically acceptable solutions and taking into account the definition of x , one gets two possible constraints:

$$\alpha \geq 1 \quad 0 < x \leq \min\left(\frac{1}{4\alpha^2}, H^2\right) \tag{17.24}$$

$$\frac{1}{2} < \alpha < 1 \quad \frac{1-\alpha}{\alpha} < H^2 \quad \frac{1-\alpha}{\alpha} < x \leq \min\left(\frac{1}{4\alpha^2}, H^2\right) . \tag{17.25}$$

In order to get a qualitative description of the λ -line, let us first notice that, according to (17.22), y diverges for $x \rightarrow 0$ and $x \rightarrow (1-\alpha)/\alpha$, i.e., at the lower boundary of the allowed interval of x according to (17.24) and (17.25) respectively. Equation (17.21) shows that the behavior of q is different in the two cases: in the former case q diverges as well, whereas in the latter it tends to a finite value q_{\max} given by

$$q_{\max} = \sqrt{\frac{1}{2\alpha-1} \ln\left(\frac{\alpha H^2}{1-\alpha}\right)} . \tag{17.26}$$

However, according to (17.23) both cases correspond to $z \rightarrow \pm 1$. Hence, as for the spinodal, the density along the λ -line diverges for $z \rightarrow \pm 1$. Moreover, from (17.22), (17.23) we get at once

$$\frac{dz}{dx} = \mp \frac{2\alpha}{\sqrt{1-4\alpha^2x}} \frac{2\alpha^2x + \alpha - 1}{(1-2\alpha x)^2} \tag{17.27}$$

$$\frac{dy}{dz} = \frac{dy}{dx} \frac{dx}{dz} = \pm \frac{(1-x)(1-2\alpha x)^2 \sqrt{1-4\alpha^2x}}{2(2\alpha-1)x(\alpha x + \alpha - 1)^2} \left(\frac{H}{x^\alpha}\right)^{1/(2\alpha-1)} , \tag{17.28}$$

where the upper and lower signs refer to $z > 0$ and $z < 0$ respectively. For $z \geq 0$, (17.28) shows that y is an increasing function of z , whereas (17.27) shows that z is a decreasing function of x . Therefore, the minimum of y is obtained for the minimum of z , which in turn corresponds to the maximum of x . According to (17.24), (17.25), two situations may occur:

If $H \geq \frac{1}{2\alpha}$, then $x_{\max} = \frac{1}{4\alpha^2}$, $z_{\min} = 0$, $y_{\min} = \frac{4\alpha}{2\alpha-1} (2\alpha H)^{1/(2\alpha-1)}$.

If $H < \frac{1}{2\alpha}$, then $x_{\max} = H^2$, $z_{\min} = \frac{\sqrt{1-4\alpha^2H^2}}{1-2\alpha H^2}$, $y_{\min} = \frac{1-2\alpha H^2}{H(\alpha H^2 + \alpha - 1)}$.

In the former case, z spans the whole concentration axis $-1 < z < 1$ and the density along the λ -line attains its minimum at the equimolar concentration $z = 0$, just like the spinodal. In the latter case, instead, the λ -line is found only outside the interval $-z_{\min} < z < z_{\min}$. Moreover, for $z = \pm z_{\min}$ one has $q = 0$, so that z_{\min} must belong to the spinodal as well. Therefore, at $z = \pm z_{\min}$ the λ -line bifurcates from the spinodal.

To complete the picture, one has to establish how the λ -line is located with respect to the spinodal in the cases in which both of them are present. Let us assume that either (17.24) or (17.25) is verified, so that the λ -line exists.

If $H \geq 1$, the spinodal does not exist. Moreover, (17.24), (17.25) show that, if a λ -line exists, then it must be $2\alpha > 1$. Therefore, one has a fortiori $2\alpha H > 1$, so that the λ -line spans the whole interval $-1 < z < 1$, assuming its minimum at $z = 0$.

If, instead, $1/(2\alpha) \leq H < 1$, the spinodal exists, and both the spinodal and the λ -line span the interval $-1 < z < 1$ and assume their minima at $z = 0$. The condition that the minimum of the λ -line lies below that of the spinodal is

$$\frac{4\alpha}{2\alpha - 1} (2\alpha H)^{1/(2\alpha-1)} \leq \frac{2}{1 - H} \quad (17.29)$$

which can be rewritten as

$$(2\alpha - 1)\zeta^{-1/(2\alpha-1)} \geq 2\alpha - \zeta, \quad (17.30)$$

where we have introduced the variable $\zeta = 2\alpha H$. It is readily found that (17.30) is verified in the whole interval of allowed values of ζ , $1 \leq \zeta < 2\alpha$, the equality holding only for $\zeta = 1$. Hence, the minimum of the λ -line is indeed smaller than that of the spinodal, save for the marginal case $H = 1/(2\alpha)$, when the two minima coincide. For $z > 0$, one has $x < 1/(4\alpha^2) \leq H^2$ along the λ -line, so that according to (17.21) q never vanishes, i.e., the λ -line and the spinodal never intersect. Therefore, the λ -line always lie below the spinodal.

Finally, for $H < 1/(2\alpha)$ the spinodal exists and spans the interval $-1 < z < 1$, whereas the λ -line bifurcates from the spinodal at $z = \pm z_{\min}$. It is also found that dy/dz at $z = \pm z_{\min}$ assumes the same value on both curves, so that the λ -line and the spinodal are tangent at the point of bifurcation. Let us then compare them for $z \rightarrow \pm 1$, i.e., for x approaching its lower bound as given by either (17.24) or (17.25). If (17.18) for y on the spinodal is evaluated at the z of (17.23), we find that the condition that the λ -line lies below the spinodal is equivalent to

$$2\alpha(1 - H^2)(x^{\alpha-1} H)^{1/(2\alpha-1)} < H(1 - 2\alpha x) + \sqrt{4\alpha x(\alpha x + \alpha - 1) + H^2(1 - 4\alpha^2 x)}. \quad (17.31)$$

For $\alpha \geq 1$, the lower bound of x is $x = 0$, for which (17.31) is trivially satisfied as it yields $H > 0$ for $\alpha > 1$ and $H^2 > 0$ for $\alpha = 1$. For $1/2 < \alpha < 1$, the lower bound of x is $x = (1 - \alpha)/\alpha$. If we set $\gamma = (1 - \alpha)/\alpha$, $\eta = H^2/\gamma$, (17.31) can be rewritten as

$$1 - \gamma\eta < (1 - \gamma)\eta^{-\gamma/(1-\gamma)}, \quad (17.32)$$

which is satisfied in the whole interval of allowed values of η , $1 < \eta < (1 + \gamma)^2/(4\gamma)$. Therefore, for $z \rightarrow \pm 1$ the λ -line always lies below the spinodal. Since $z = \pm z_{\min}$ are the only points on the λ -line such that $q = 0$, the spinodal and the λ -line never intersect for $z > z_{\min}$, so that the λ -line lies below the spinodal in the whole interval $z_{\min} < z < 1$ as well as, of course, in the symmetric interval $-1 < z < -z_{\min}$.

17.3 Summary

We have investigated the phase behavior of athermal, binary mixtures such that the interaction between particles of species i and particles of species j is a Gaussian with height ϵ_{ij} and width R_{ij} . Specifically, we have considered the symmetric case $\epsilon_{11} = \epsilon_{22}$, $R_{11} = R_{22}$, and determined which conditions cause an instability of the homogeneous, mixed phase, leading to either inhomogeneous structures, or a demixing transition into two bulk phases. These occurrences are signaled respectively by the occurrence of the λ - and spinodal lines, which mark the boundaries of the stability domain of the homogeneous phase in the density-concentration plane. We found that the behavior of the system is determined by the interaction parameters ϵ_{ij} , R_{ij} via the positive quantities α and H defined in (17.15), (17.16). Our results are summarized as follows:

- The λ -line exists for

$$\alpha \geq 1, \text{ any } H$$

$$\frac{1}{2} < \alpha < 1, \frac{1 - \alpha}{\alpha} < H^2.$$

In both cases, the density and the wave vector of the instability increase as one moves along the λ -line towards the boundaries of the concentration interval. At the boundaries, the density always diverges, whereas the wave vector diverges in the former case, and tends to a finite value in the latter.

- If either of the above conditions is satisfied so that the λ -line does exist, three scenarios may occur:
 1. $H \geq 1$: The spinodal does not exist, and the λ -line spans the whole concentration interval, achieving its minimum at equimolar concentration.
 2. $1/(2\alpha) \leq H < 1$: Both the λ and the spinodal lines span the whole concentration interval and have their minima at equimolar concentration, but the λ -line lies below the spinodal.
 3. $H < 1/(2\alpha)$: The spinodal spans the whole concentration interval, whereas the λ -line is found only at suitably high or low concentration, and bifurcates from the spinodal at the boundaries of the forbidden interval. In the allowed domain, the λ -line lies again below the spinodal.

Figures 17.1, 17.2 and 17.3 display the λ -line given by (17.22), (17.23) and the spinodal (when existent) given by (17.18) in terms of the total density ρ and concentration ϕ for three choices of interaction parameters corresponding to the above scenarios.

On the basis of the present analysis, $\alpha > 1/2$ is a *necessary* condition for the occurrence of the λ -line in symmetric mixtures. In terms of the interaction ranges R_{ij} this means $R_{12} < R_{11} = (R_{11} + R_{22})/2$, i.e., negative non-additivity. Also, a *sufficient* condition is given by $\alpha > 1$, i.e., $R_{11} > \sqrt{2}R_{12}$. Indeed, in a former study [11] this sufficient condition was shown to hold more generally even for non-symmetric

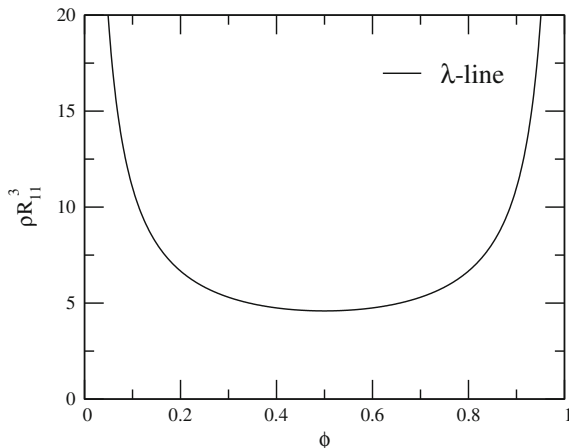


Fig. 17.1 λ -line in the density-concentration plane for $\epsilon_{12}/\epsilon_{11} = 1.5$, $R_{12}/R_{11} = 0.8$, corresponding to $H > 1$ (see text)

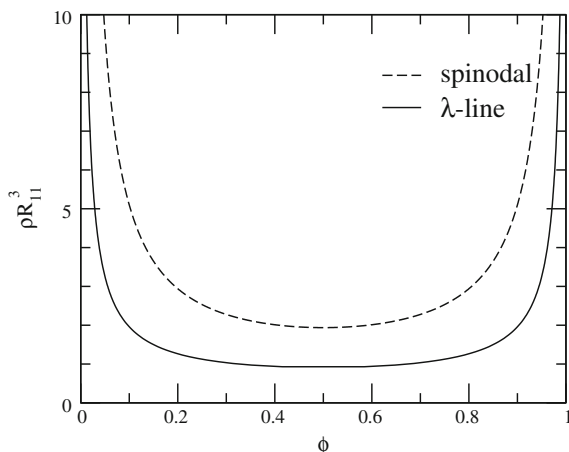


Fig. 17.2 λ -line and spinodal in the density-concentration plane for $\epsilon_{12}/\epsilon_{11} = 4$, $R_{12}/R_{11} = 2/3$, corresponding to $1/(2\alpha) < H < 1$ (see text)

mixtures. Conversely, in the non-symmetric case negative non-additivity is not a necessary condition, as strongly asymmetric mixtures with very different values of R_{11} and R_{22} may satisfy the additivity condition $R_{12} = (R_{11} + R_{22})/2$ and yet have $R_{11} > \sqrt{2}R_{12}$.

Establishing the conditions for the formation of inhomogeneous phases does not say anything on which kind of phases will actually populate the phase diagram of a specific mixture. This would require a much deeper investigation based on methods such as density-functional theory. To our knowledge, this study has been carried out

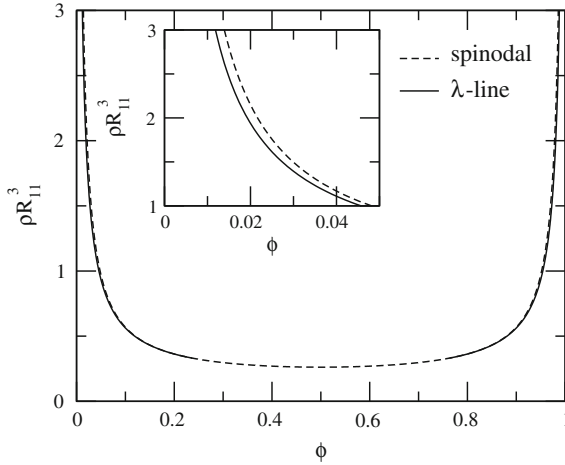


Fig. 17.3 λ -line and spinodal in the density-concentration plane for $\epsilon_{12}/\epsilon_{11}=8$, $R_{12}/R_{11}=2/3$, corresponding to $H < 1/(2\alpha)$ (see text). The inset is an enlargement of the low-concentration region

so far only for a specific instance of a non-symmetric mixture [10, 11] which does not have a spinodal line, similarly to case 1 discussed above. In the future, it may be interesting to perform a similar study also for systems belonging to cases 2 or 3, such that both microphase formation and a demixing transition can take place. The present results show that the spinodal is always preempted by the λ -line at all concentrations at which both kind of instabilities occur, even though the two may be very close to each other, as in Fig. 17.3. However, one has to take into account that, even though those loci give an overall idea of the shape of the phase diagram, they do not generally coincide with the actual first-order transition lines. In order to establish precisely which transitions are observed and where, the free energies of all the phases involved should be compared. Again, this requires going beyond the analysis developed here.

References

1. P.J. Flory, J. Krigbaum, Statistical mechanics of dilute polymer solutions II. *J. Chem. Phys.* **18**, 1086 (1950). <https://doi.org/10.1063/1.1747866>
2. J. Dautenhahn, C.K. Hall, Monte Carlo simulation of off-lattice polymer chains: effective pair potentials in dilute solution. *Macromolecules* **27**, 5399 (1994). <https://doi.org/10.1021/ma00097a021>
3. A.A. Louis, P.G. Bolhuis, J.P. Hansen, E.J. Meijer, Can polymer coils be modeled as “soft colloids”? *Phys. Rev. Lett.* **85**, 2522 (2000). <https://doi.org/10.1103/PhysRevLett.85.2522>
4. A.A. Louis, P.G. Bolhuis, J.-P. Hansen, Mean-field fluid behavior of the Gaussian core model. *Phys. Rev. E* **62**, 7961 (2000). <https://doi.org/10.1103/PhysRevE.62.7961>
5. R. Finken, J.-P. Hansen, A.A. Louis, Phase separation of penetrable core mixtures. *J. Stat. Phys.* **110**, 1015 (2003). <https://doi.org/10.1023/A:1022136624854>

6. A.J. Archer, R. Evans, Binary Gaussian core model: fluid-fluid phase separation and interfacial properties. *Phys. Rev. E* **64**, 041501 (2001). <https://doi.org/10.1103/PhysRevE.64.041501>
7. A.J. Archer, R. Evans, Wetting in the binary Gaussian core model. *J. Phys.: Condens. Matter* **14**, 1131 (2002). <https://doi.org/10.1088/0953-8984/14/6/302>
8. A.J. Archer, R. Evans, Solvent-mediated interactions and solvation close to fluid-fluid phase separation: a density functional treatment. *J. Chem. Phys.* **118**, 9726 (2003). <https://doi.org/10.1063/1.1570406>
9. I.O. Götze, A.J. Archer, C.N. Likos, Structure, phase behavior, and inhomogeneous fluid properties of binary dendrimer mixtures. *J. Chem. Phys.* **124**, 084901 (2006). <https://doi.org/10.1063/1.2172596>
10. A.J. Archer, C.N. Likos, R. Evans, Soft-core binary fluid exhibiting a λ -line and freezing to a highly delocalized crystal. *J. Phys.: Condens. Matter* **16**, L297 (2004). <https://doi.org/10.1088/0953-8984/16/23/L03>
11. D. Pini, A. Parola, L. Reatto, An unconstrained DFT approach to microphase formation and application to binary Gaussian mixtures. *J. Chem. Phys.* **143**, 034902 (2015). <https://doi.org/10.1063/1.4926469>
12. J.-P. Hansen, I.R. McDonald, *Theory of Simple Liquids* (Academic Press, London, 2006)

Chapter 18

Tutorial: γ_5 in Dimensional Regularization in Dimensional Regularization



Mario Raciti

Abstract Dimensional Regularization (DR) is the most powerful tool to handle renormalization in perturbative Quantum Field Theory. It allows indeed a very effective way to satisfy the symmetry requirements in gauge theory. The limitations of the method arise from the definition of γ_5 , which has been the most controversial point, since the work of 't Hooft and Veltman. After a short introduction on renormalization and the regularization problems, the definition of γ_5 in DR is considered in the light of the axial anomaly.

18.1 Introduction

Quantum field theory has to address the problem of UV-divergencies, namely the fact that perturbative expansion of Green functions turns out to be expressed by integrals which are divergent for large momenta. The problem is solved by renormalization theory [1].

The removal of infinities must be performed so that the symmetry properties of the theory are preserved. To this aim for gauge symmetry a very important tool is Dimensional Regularization (DR) [2–8] which, at least for vectorial coupling, makes the procedure to some degree automatic. The difficulties are limited to the implementation of chiral symmetries, which seem to be bound to the physical four dimensional space-time.

After a brief account of the renormalization theory in Sect. 18.2, in Sect. 18.3 we describe the DR and then in Sect. 18.4 the problem of the definition of γ_5 , the most

M. Raciti (✉)
Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano,
Milan, Italy
e-mail: mario.raciti@mi.infn.it

M. Raciti
INFN-Sezione di Milano, Milan, Italy

critical point. The discussion, as in the original paper by t’Hooft and Veltman[2], leans on the paramount example of the chiral symmetry in QED [9, 10]. The very existence of the axial anomaly puts indeed strong constraints on this definition. A last section of conclusions follows.

18.2 Renormalization and Regularization

Perturbative QFT, beyond the tree level approximation, encounters some difficulties. In momenta space the loop integrals are indeed in general divergent for large values of integration variables. For instance in $\frac{g}{4!}\phi^4$ scalar theory consider the amplitude of Fig. 18.1:

Its analytic expression is

$$\Gamma(p) = \frac{(ig)^2}{2!} \int \frac{d^4q}{(2\pi)^4} \Delta(p - q)\Delta(q), \tag{18.1}$$

where $\Delta(q) = \frac{i}{q^2 - m^2 + i\epsilon}$ is the Feynman propagator and the integral diverges logarithmically for large q .

Renormalization theory is a series of consistent iterative prescriptions which replaces divergent integrals as the former with convergent ones.

In this case the renormalized graph could be written (up to a constant) as

$$\Gamma_{\text{ren}}(p) = \frac{(ig)^2}{2!} \int \frac{d^4q}{(2\pi)^4} (\Delta(p - q)\Delta(q) - \Delta^2(q)). \tag{18.2}$$

In this example the expression of the renormalized graph has been obtained with a substitution rule on the integrand; it is possible to follow this procedure in the so-called BPHZ approach [1], but in general one prefers to modify systematically the Feynman rules in such a way to produce, at intermediate step, well behaved convergent expressions. For instance in this case we could substitute the original propagator according to $\Delta(q) \rightarrow \Delta^\Lambda(q) = \frac{i}{q^2 - m^2 + i\epsilon} - \frac{i}{q^2 - \Lambda^2 + i\epsilon}$, which as $\Lambda \rightarrow \infty$ becomes the original propagator and has a more convergent behaviour for large momenta. Using the regularized Feynman rules, at intermediate step we write for the diagram

$$\frac{(ig)^2}{2!} \int \frac{d^4q}{(2\pi)^4} (\Delta^\Lambda(q - p)\Delta^\Lambda(q) - (\Delta^\Lambda(q))^2) \tag{18.3}$$

Fig. 18.1 A four-point amplitude for a simple scalar theory



where, at finite Λ , we subtract two separately convergent integrals. The second one corresponds to add a new term in the lagrangian (called counterterm) proportional to the interaction term, to be taken into account also to next orders of the perturbative expansion. The regularized propagator Δ^Λ has the same Lorentz covariant properties of the original one, and the same holds for the regularized Green functions, as a consequence one can prove that the counterterms turn out to be Lorentz scalar polynomials in the fields, indeed a good regularization should maintain the symmetries of the formal divergent theory. Regularizations of the type just described are called Pauli-Villars [1] regularizations, they introduce a mass parameter, Λ in our example, but this breaks gauge symmetry (a gauge invariant version has however been devised for the abelian case).

18.3 Dimensional Regularization

The problem is solved by Dimensional Regularization [2–8]. The divergent or convergent character of a Feynman integral depends on the dimension D of the space-time, and for suitable low dimension, consider our simple previous example, is convergent. Moreover by almost explicit calculations the D -dependence of the Feynman integrals turns out to be analytic, and has poles as $D \rightarrow 4$. This is the intuitive idea in which D plays the role that the mass parameter Λ plays in Pauli-Villars regularization.

On the other hand the gauge symmetry can be formulated in every space-time dimension D : at least for vectorial couplings Ward identities, which control and describe the symmetry, do not involve in critical way the dimensionality of the space.

Beyond the intuitive understanding in [6–8] has been shown that DR, as a set of formal rules defining the Green's functions, is a correct regularization. What is proven is that the theory defined by DR fulfills a set of consistency requirements, called Quantum Action Principle.

Let's consider now our simple one loop diagram. Introducing Feynman parametrization we get

$$\begin{aligned}
 & \frac{(ig)^2}{2} \int \frac{d^D q}{(2\pi)^D} \frac{i}{q^2 - m^2 + i\epsilon} \frac{i}{(q-p)^2 - m^2 + i\epsilon} = \\
 & = \frac{g^2}{2} \int \frac{d^D q}{(2\pi)^D} \int_0^1 dx \frac{1}{((q^2 - m^2 + i\epsilon)x + (1-x)((q-p)^2 - m^2 + i\epsilon))^2} = \\
 & = \frac{g^2}{2} \int_0^1 dx \int \frac{d^D q}{(2\pi)^D} \frac{1}{(q^2 + x(1-x)p^2 - m^2 + i\epsilon)^2} = \\
 & = i \frac{g^2}{4} \frac{\Gamma(\frac{4-D}{2})}{(4\pi)^{D/2}} \int_0^1 dx (m^2 - x(1-x)p^2)^{\frac{4-D}{2}}.
 \end{aligned} \tag{18.4}$$

In the third line we performed a translation in the momentum integration variable $q \rightarrow q + (1-x)p$, and in the last, after an anticlockwise rotation in the complex

plane for the dq_0 integration, we used a standard integral formula; it is in this last equality that the D -dependence becomes analytic. Expanding (18.4) in Laurent series for $D \rightarrow 4$ we get a pole term, which must be subtracted away, and this defines the counterterm (in the Minimal Subtraction), and an order-zero term, which defines the renormalized graph value. This procedure can be generalized in such a way to deal with the renormalization of every bosonic theory, in order to include the fermions some more definitions are needed.

18.3.1 Extension to Fermions

The following rules define the D -dimensional gamma algebra and the traces properties. The starting point is

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}. \quad (18.5)$$

If S is an odd string of γ 's

$$tr[S] = 0. \quad (18.6)$$

Moreover

$$tr[I] = 4. \quad (18.7)$$

The vectorial indexes, μ, ν etc. run in a D -dimensional space (as in the bosonic case), therefore $g_\mu^\mu = D$. For D even the two first equations could be satisfied by $2^{D/2}$ -dimensional matrices, in any case (18.5)–(18.7) are a set of consistent formal rules (the trace is by definition cyclic) correct for $D \rightarrow 4$ [2, 6–8]. Starting from (18.5) one can compute the D -dimensional trace of a string of γ matrices, in similar way to the familiar $D = 4$ case, see Appendix C of [2].

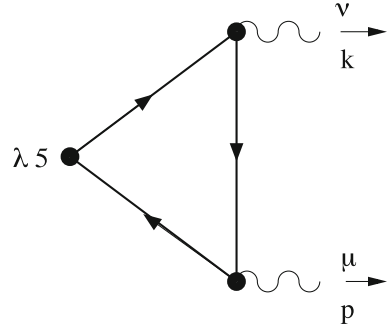
The rules of this section allow a powerful and elegant renormalization of vectorial gauge theory, but we have still to consider the definition of γ_5 . It is a crucial point in order to define chiral gauge theories.

18.4 ABJ Anomaly and the Definition of γ_5

A good point to tackle the γ_5 problem in DR is the computation of the anomaly coefficient in DR. Consider the following transformation of the fermionic field in spinorial QED:

$$\psi \rightarrow \exp(i\alpha\gamma_5)\psi \quad (18.8)$$

Fig. 18.2 The basic triangle for the axial anomaly



and correspondingly

$$\bar{\psi} \rightarrow \bar{\psi} \exp(i\alpha\gamma_5). \tag{18.9}$$

A naive use of Noether theorem would lead to the equation $\partial_\lambda \bar{\psi} \gamma^\lambda \gamma_5 \psi = 2im \bar{\psi} \gamma_5 \psi$, indeed the mass term is the only not invariant term in the lagrangian under the transformations of (18.8) and (18.9), and therefore for the massless theory we should have a conserved current.

Noether theorem is a classical theorem, it does not take into account that it is by no means obvious to define product of fields at the same space-time point, namely that there are the ultraviolet divergences. Indeed the equation has to be corrected [9, 10] as

$$\partial_\lambda \bar{\psi} \gamma^\lambda \gamma_5 \psi = 2im \bar{\psi} \gamma_5 \psi + \frac{\alpha}{4\pi} \epsilon_{\mu\nu\rho\sigma} F^{\mu\nu} F^{\rho\sigma}. \tag{18.10}$$

To get the coefficient $\frac{\alpha}{4\pi}$ we take the divergence of a suitable 1PI-Green's function, $\partial_{x^\lambda} \langle 0 | T J_5^\lambda(x) A^\mu(y) A^\nu(z) | 0 \rangle_{1PI}$; assuming the fermion massless the result should be zero if the naive Noether theorem was correct (there are not “contact terms”). To leading order and in momenta space we have two Feynman diagrams, the first graph is in Fig. 18.2 the second one is the crossed ($\mu, p \leftrightarrow \nu, k$) partner.

The first aim is to show that (18.5)–(18.7) uniquely determine the anomaly of (18.10). The diagram has the following analytical expression:

$$\Gamma_5^{\lambda\mu\nu}(p, k) = -i^3 (ie)^2 \int_q \frac{\text{tr}[\gamma^\lambda \gamma_5 (\not{q} - \not{p}) \gamma^\mu \not{q} \gamma^\nu (\not{q} + \not{k})]}{(q+k)^2 (q-p)^2 q^2}. \tag{18.11}$$

In the following we use these shorthand notations: $\int_{xy} \dots = \int_0^1 dx \int_0^{1-x} dy \dots$, $\int_q \dots = \int \frac{d^D q}{(2\pi)^D} \dots$ and $\epsilon = 4 - D$.

By introducing Feynman parametrization $\Gamma_5^{\lambda\mu\nu}(p, k)$ can be transformed into

$$\Gamma_5^{\lambda\mu\nu}(p, k) = -2ie^2 \int_{xy} \int_q \frac{\text{tr}[\gamma^\lambda \gamma_5 (\not{q} - \not{p}) \gamma^\mu \not{q} \gamma^\nu (\not{q} + \not{k})]}{((q + yk - px)^2 - (yk - px)^2 + xp^2 + yk^2)^3} \quad (18.12)$$

and performing a translation in the momentum variable into

$$\Gamma_5^{\lambda\mu\nu}(p, k) = -2ie^2 \int_{xy} \int_q \frac{\text{tr}[(\not{q} + \not{r} - \not{p}) \gamma^\mu ((\not{q} + \not{r}) \gamma^\nu ((\not{q} + \not{r} + \not{k}) \gamma^\lambda \gamma_5)]}{(q^2 + 2k \cdot pxy + k^2y(1-y) + p^2x(1-x))^3}, \quad (18.13)$$

where $r = xp - yk$.

Being now the denominator function of q^2 , only the even powers of q give a nonvanishing contribution to the integral, and using the Lorentz D -dimensional invariance $q^\alpha q^\beta$ can be replaced by $g^{\alpha\beta} q^2/D$. Therefore the new numerator is

$$\begin{aligned} & \frac{q^2}{D} \left(\text{tr}[\gamma^\alpha \gamma^\mu \gamma_\alpha \gamma^\nu (\not{r} + \not{k}) \gamma^\lambda \gamma_5] + \text{tr}[\gamma^\alpha \gamma^\mu \not{r} \gamma^\nu \gamma_\alpha \gamma^\lambda \gamma_5] + \right. \\ & \left. + \text{tr}[(\not{r} - \not{p}) \gamma^\mu \gamma^\alpha \gamma^\nu \gamma_\alpha \gamma^\lambda \gamma_5] \right) + \text{tr}[(\not{r} - \not{p}) \gamma^\mu \not{r} \gamma^\nu (\not{r} + \not{k}) \gamma^\lambda \gamma_5]. \end{aligned}$$

The three addenda in q^2 lead to a divergent integral by power counting, while the last term gives a finite contribution, correspondingly we will separate $\Gamma_5^{\lambda\mu\nu}(p, k)$ as $\Gamma_5^{\lambda\mu\nu}(p, k) = \Gamma_{5\text{div}}^{\lambda\mu\nu}(p, k) + \Gamma_{5\text{fin}}^{\lambda\mu\nu}(p, k)$, moreover for the divergent terms, using (18.5) and the cyclicity of the trace, we can put the momenta r, k and p in a standard position along the trace, arriving at

$$\begin{aligned} & \frac{q^2}{D} \left(-2\text{tr}[\gamma^\mu \gamma^\nu (3\not{r} + \not{k} - \not{p}) \gamma^\lambda \gamma_5] + \epsilon \text{tr}[\gamma^\mu \gamma^\nu (\not{r} + \not{k} - \not{p}) \gamma^\lambda \gamma_5] + \right. \\ & + 4g^{\mu\nu} \text{tr}[\not{r} \gamma^\lambda \gamma_5] - 4(2r - p)^\mu \text{tr}[\gamma^\nu \gamma^\lambda \gamma_5] - 4(p - r)^\nu \text{tr}[\gamma^\mu \gamma^\lambda \gamma_5] + \\ & \left. + 2\epsilon(p^\nu \text{tr}[\gamma^\mu \gamma^\lambda \gamma_5] + (r - p)^\mu \text{tr}[\gamma^\nu \gamma^\lambda \gamma_5]) \right). \end{aligned}$$

Omitting for the moment the traces with only two gamma matrices, that we will reconsider next, we can write:

$$\begin{aligned} & \Gamma_{5\text{div}}^{\lambda\mu\nu}(p, k) = \\ & = \frac{-2ie^2}{D} \int_{xy} (-2\text{tr}[\gamma^\mu \gamma^\nu (3\not{r} + \not{k} - \not{p}) \gamma^\lambda \gamma_5] + \epsilon \text{tr}[\gamma^\mu \gamma^\nu (\not{r} + \not{k} - \not{p}) \gamma^\lambda \gamma_5]) \cdot \\ & \cdot \int_q \frac{q^2}{(q^2 + 2k \cdot pxy + k^2y(1-y) + p^2x(1-x))^3}. \end{aligned} \quad (18.14)$$

The inner integral can be evaluated with the Wick rotation, it gives

$$\begin{aligned}
& \int_q \frac{q^2}{(q^2 + 2k \cdot pxy + k^2y(1-y) + p^2x(1-x))^3} = \\
& = \frac{i}{(4\pi)^{D/2}} \frac{D}{4} \Gamma\left(\frac{\epsilon}{2}\right) (-2k \cdot pxy - k^2y(1-y) - p^2x(1-x))^{-\frac{\epsilon}{2}} = \\
& = \frac{i}{(4\pi)^{D/2}} \frac{D}{4} \Gamma\left(\frac{\epsilon}{2}\right) \cdot \\
& \cdot \left(1 - \frac{\epsilon}{2} \log((-2k \cdot pxy - k^2y(1-y) - p^2x(1-x)) + \dots)\right).
\end{aligned} \tag{18.15}$$

The gamma function has a first order pole in $\epsilon = 0$ but its residue in (18.14) is proportional to $\int_{xy} (-2\text{tr}[\gamma^\mu \gamma^\nu (3\not{f} + \not{k} - \not{p}) \gamma^\lambda \gamma_5])$, and because of the integration in the Feynman parameters it actually vanishes, thus the result is finite and we have

$$\begin{aligned}
\Gamma_{5\text{div}}^{\lambda\mu\nu}(p, k) &= \frac{e^2}{2} \frac{1}{(4\pi)^{D/2}} \Gamma\left(\frac{\epsilon}{2}\right) \epsilon \int_{xy} \left(\text{tr}[\gamma^\mu \gamma^\nu (\not{f} + \not{k} - \not{p}) \gamma^\lambda \gamma_5] + \right. \\
&+ \left. \text{tr}[\gamma^\mu \gamma^\nu (3\not{f} + \not{k} - \not{p}) \gamma^\lambda \gamma_5] \cdot \log((-2k \cdot pxy - k^2y(1-y) - p^2x(1-x))) \right) \\
&\rightarrow \frac{e^2}{(4\pi)^2} \frac{1}{3} \text{tr}[\gamma^\mu \gamma^\nu (\not{k} - \not{p}) \gamma^\lambda \gamma_5] + \frac{e^2}{(4\pi)^2} \int_{xy} \text{tr}[\gamma^\mu \gamma^\nu (3\not{f} + \not{k} - \not{p}) \gamma^\lambda \gamma_5] \cdot \\
&\cdot \log(-2k \cdot pxy - k^2y(1-y) - p^2x(1-x)).
\end{aligned} \tag{18.16}$$

The traces in this last formula multiply finite quantities for $D \rightarrow 4$, so the only property tested, is that “in same way” the traces assume their value when $D = 4$, but no extrapolation to $D \neq 4$ is actually required. In similar way if we reconsider the terms depending on the traces with only two gamma matrices, and sum the corresponding quantities of the crossed diagram, it turns out that also these traces are multiplied by convergent integrals, so the final result depends only on their values in $D = 4$, which is zero, and thus their contribute vanishes.

In $\Gamma_{5\text{fin}}^{\lambda\mu\nu}(p, k)$ the traces are computed in $D = 4$, the starting point is

$$\Gamma_{5\text{fin}}^{\lambda\mu\nu}(p, k) = \frac{e^2}{(4\pi)^2} \int_{xy} \frac{\text{tr}[(\not{f} - \not{p}) \gamma^\mu \not{f} \gamma^\nu (\not{f} + \not{k}) \gamma^\lambda \gamma_5]}{2k \cdot pxy + k^2y(1-y) + p^2x(1-x)}, \tag{18.17}$$

and we arrive at

$$\begin{aligned}
(p+k)_\lambda \Gamma_{5\text{fin}}^{\lambda\mu\nu}(p, k) &= \frac{e^2}{(4\pi)^2} \text{tr}[\gamma^\mu \gamma^\nu \not{k} \not{p} \gamma_5] \cdot \\
&\cdot \int_{xy} \frac{((1-x)^2x + x^2y)p^2 + (y(1-y)^2 + y^2x)k^2 + 2k \cdot pxy(2-x-y)}{2k \cdot pxy + k^2y(1-y) + p^2x(1-x)}.
\end{aligned} \tag{18.18}$$

The corresponding quantity arising from $\Gamma_{5\text{div}}^{\lambda\mu\nu}(p, k)$ is:

$$\begin{aligned}
 (p+k)_\lambda \Gamma_{5 \text{ div}}^{\lambda\mu\nu}(p,k) &= \frac{e^2}{(4\pi)^2} \text{tr}[\gamma^\mu \gamma^\nu \not{k} \not{p} \gamma_5] \cdot \\
 &\cdot \left(\frac{2}{3} + \int_{xy} (2 - 3(x+y)) \log(-2k \cdot pxy - k^2y(1-y) - p^2x(1-x)) \right) = \\
 &= \frac{e^2}{(4\pi)^2} \text{tr}[\gamma^\mu \gamma^\nu \not{k} \not{p} \gamma_5] \cdot \\
 &\cdot \left(\frac{2}{3} - \int_{xy} (1-x-y) \frac{k^2y(1-2y) + p^2x(1-2x) + 4k \cdot pxy}{2k \cdot pxy + k^2y(1-y) + p^2x(1-x)} \right).
 \end{aligned} \tag{18.19}$$

The expression of the last line has been obtained with an integration by parts. Indeed if $\lambda = x + y$ and $\alpha = \frac{x}{x+y}$, we can write

$$\begin{aligned}
 &\int_{xy} (2 - 3\lambda) \log(f(\lambda, \alpha)) = \\
 &= \int_0^1 d\alpha \int_0^1 d\lambda \lambda (2 - 3\lambda) \log(f(\lambda, \alpha)) = \\
 &= \int_0^1 d\alpha \int_0^1 d\lambda \frac{d(\lambda^2 - \lambda^3)}{d\lambda} \log(f(\lambda, \alpha)) = \\
 &= - \int_0^1 d\alpha \int_0^1 d\lambda (\lambda^2 - \lambda^3) \frac{\partial \log(f(\lambda, \alpha))}{\partial \lambda}.
 \end{aligned}$$

In terms of the original variables

$$\int_{xy} (2 - 3(x+y)) \log(f(x, y)) = - \int_{xy} (1-x-y) \frac{\lambda \partial \log(f(\lambda, \alpha))}{\partial \lambda} \Big|_{\lambda=x+y; \alpha=\frac{x}{x+y}},$$

which gives (18.19), then

$$(p+k)_\lambda (\Gamma_{5 \text{ div}}^{\lambda\mu\nu}(p,k) + \Gamma_{5 \text{ fin}}^{\lambda\mu\nu}(p,k)) = \frac{e^2}{(4\pi)^2} \left(\frac{2}{3} + \int_{xy} (x+y) \right) = \frac{e^2}{(4\pi)^2}. \tag{18.20}$$

The crossed graph yields exactly the same contribution, thus the final coefficient is $\frac{2e^2}{(4\pi)^2}$, which is the well known expression of the anomaly. We stress that we did not use any algebra of γ_5 , and actually only minimal hypotheses on the traces: in DR the anomaly is a consequence of the (18.5)–(18.7).

This “minimal” computation of the anomaly can be extended to graph including more γ_5 in non abelian theory [11]. On the other hand one can undertake the usual combinatorial proof of Ward identities:

$$\begin{aligned}
(p+k)_\lambda \Gamma_5^{\lambda\mu\nu}(p,k) &= -ie^2 \int_q \text{tr} \left[\frac{1}{q+k} (\not{k} + \not{p}) \gamma_5 \frac{1}{q-p} \gamma^\mu \frac{1}{q} \gamma^\nu \right] = \\
&= -ie^2 \int_q \text{tr} \left[\frac{1}{q+k} ((q+k) - (q-p)) \gamma_5 \frac{1}{q-p} \gamma^\mu \frac{1}{q} \gamma^\nu \right],
\end{aligned} \tag{18.21}$$

using

$$\frac{1}{q+k} ((q+k) - (q-p)) \gamma_5 \frac{1}{q-p} = \gamma_5 \frac{1}{q-p} + \frac{1}{q+k} \gamma_5 - \frac{1}{q+k} \{(q-p), \gamma_5\} \frac{1}{q-p}$$

we arrive at

$$\begin{aligned}
(p+k)_\lambda \Gamma_5^{\lambda\mu\nu}(p,k) &= -ie^2 \int_q \text{tr} \left[\gamma_5 \frac{1}{q-p} \gamma^\mu \frac{1}{q} \gamma^\nu \right] + \\
&\quad -ie^2 \int_q \text{tr} \left[\frac{1}{q+k} \gamma_5 \gamma^\mu \frac{1}{q} \gamma^\nu \right] + \\
&\quad -ie^2 \int_q \text{tr} \left[\frac{1}{q+k} \{(p-q), \gamma_5\} \frac{1}{q-p} \gamma^\mu \frac{1}{q} \gamma^\nu \right].
\end{aligned} \tag{18.22}$$

The previous equations are correct in DR, the sum of the first two addenda and the corresponding terms of the partner graph turns out to be dependent only on $\text{tr}[\gamma_5]$: the anomaly arises from the last integral of (18.22).

The conclusion is that in DR $\{\gamma^\mu, \gamma_5\} = 0$, the so-called Naive Dimensional Regularization (NDR), is in general wrong, for at least some μ .

18.5 The 't HV Solution

In their paper [2] 't Hooft and Veltman give the following solution to the problem. γ_5 is defined as in $D = 4$

$$\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3, \tag{18.23}$$

and the following (anti)commutation rules then follow:

$$\{\gamma^\mu, \gamma_5\} = 0 \quad \mu = 0, \dots, 3. \tag{18.24}$$

$$[\gamma^\mu, \gamma_5] = 0 \quad \mu > 3. \tag{18.25}$$

Before the analytic continuation $D > 4$ is understood, but the external momenta, p and k in (18.22), are taken physical, in the sense that only their first four components are different from zero, only the momenta running in the loop integrals are actually continued.

Because of definition (18.23) the trace of strings of γ matrices including γ_5 are consequences of (18.5), in particular $\text{tr}[\gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma \gamma_5]$ vanish if one of the vectorial indexes is nonphysical, otherwise is given by the usual four dimensional tensor $-4i\epsilon^{\mu\nu\rho\sigma}$ (see the Appendix of [2] for other useful formulas).

For the sake of generality and in order to avoid the specification of the value of indexes, (18.23)–(18.25) are often formulated by means of two kind of indexes [6–8], it is however useful to introduce the following notation: a ‘ $\hat{}$ ’ will denote the operation of projection in the nonphysical directions, so we can write $\gamma_{\hat{\rho}}$, \hat{q} and we have $\hat{p} = \hat{k} = 0$.

The first two terms in r.h.s of (18.22) vanish, being p physical for the last one we have

$$\begin{aligned} & -ie^2 \int_q \text{tr} \left[\frac{1}{q+k} \{(\not{p}-\not{q}), \gamma_5\} \frac{1}{q-\not{p}} \gamma^\mu \frac{1}{q} \gamma^\nu \right] = \\ & = 2ie^2 \int_q \frac{\text{tr}[(q-\not{p})\gamma^\mu \not{q} \gamma^\nu (q+\not{k})\hat{q} \gamma_5]}{(q-p)^2 q^2 (q+k)^2}. \end{aligned} \quad (18.26)$$

Introducing the Feynman parametrization and performing the usual translation we can then write

$$\begin{aligned} & (p+k)_\lambda \Gamma_5^{\lambda\mu\nu}(p, k) = \\ & = 4ie^2 \int_{xy} \int_q \frac{\text{tr}[(q-\not{p}+\not{t})\gamma^\mu (q+\not{t})\gamma^\nu (q+\not{k}+\not{t})\hat{q} \gamma_5]}{(q^2 + 2k \cdot pxy + k^2y(1-y) + p^2x(1-x))^3}. \end{aligned} \quad (18.27)$$

Only the terms quadratic in the loop momentum give a not-vanishing contribution to the numerator; using the replacement $q^\rho q^{\hat{\sigma}} \rightarrow g^{\hat{\rho}\hat{\sigma}} \frac{q^2}{D}$ and the (anti)commutation rules (18.24) and (18.25) it turns out to be

$$\begin{aligned} & \frac{q^2}{D} \left(-\epsilon(\text{tr}[\gamma^\mu \not{t} \gamma^\nu (\not{t}+\not{k}) \gamma_5] + \text{tr}[(\not{t}-\not{p})\gamma^\mu \not{t} \gamma^\nu \gamma_5] + \right. \\ & + \text{tr}[(\not{t}-\not{p})\gamma^\mu \gamma^\nu ((\not{t}+\not{k}) \gamma_5]) + \\ & + 2\text{tr}[(\not{t}-\not{p})\gamma^\mu (\not{t}+\not{k})\gamma^{\hat{\nu}} \gamma_5] - 2\text{tr}[(\not{t}-\not{p})\gamma^\mu \gamma^\nu (\hat{t}+\hat{k}) \gamma_5] \left. \right) = \\ & - \frac{q^2}{D} \epsilon \text{tr}[\gamma^\mu \gamma^\nu \not{k} \not{p} \gamma_5]. \end{aligned}$$

The traces in the third line disappeared because they contain at least one non physical index. Then

$$\begin{aligned}
(p+k)_\lambda \Gamma_5^{\lambda\mu\nu}(p,k) &= -4ie^2 \frac{\epsilon}{D} \text{tr}[\gamma^\mu \gamma^\nu \not{k} \not{p} \gamma_5] \cdot \\
&\cdot \int_{xy} \int_q \frac{q^2}{(q^2 + 2k \cdot pxy + k^2 y(1-y) + p^2 x(1-x))^3} = \\
&= -4ie^2 \frac{\epsilon}{D} \text{tr}[\gamma^\mu \gamma^\nu \not{k} \not{p} \gamma_5] \int_{xy} \left(\frac{i}{(4\pi)^2} \frac{2}{\epsilon} + \dots \right) \\
&\rightarrow \frac{1}{(4\pi)^2} \text{tr}[\gamma^\mu \gamma^\nu \not{k} \not{p} \gamma_5].
\end{aligned} \tag{18.28}$$

Taking into account of the identical contribute of the other diagram we arrive at the anomaly again.

Equation (18.28) rests on (18.23), and its properties (18.24) and (18.25), which are in any case compatible with the more general first calculation, and they are consistent because (18.5)–(18.7) are consistent. So we are granted that different manipulations lead to the same result, as we have just checked. That this definition of γ_5 is consistent from a physical point of view, namely that the Quantum Action Principle is fulfilled, is proven in [6–8].

The axial anomaly could be also computed with a suitable Pauli-Villars regularization, where γ_5 is the usual four dimensional matrix. Now the mass parameter Λ breaks the chiral symmetry of (18.8)–(18.9), and the last addendum in (18.22) is replaced by a Λ -dependent term, from this term the anomaly arises as $\Lambda \rightarrow \infty$, see Sect. 11.5 of [12] for details.

A last comment is in order: the anomaly computed in (18.20) could be cancelled by a suitable finite counterterm. To this aim we should add a counterterm proportional to $\text{tr}[\gamma_5 \gamma^\lambda \gamma^\mu \gamma^\nu \gamma^\rho](k-p)_\rho$ to the original $\Gamma_5^{\lambda\mu\nu}(p,k)$, but this possibility is ruled out by the gauge invariance requirement, indeed being DR a gauge invariant regularization $p_\mu \Gamma_5^{\lambda\mu\nu}(p,k) = 0$ holds, and no longer the new vertex would fulfill the same equation. This impossibility is actually the essence of the anomaly.

One could think that the very fact that we computed the anomaly in the first calculation with almost no assumption on γ_5 , shows that its algebra is a fake problem, or is at least unimportant.

The same computation, at two loop level, shows that the subject is more complex. For the anomaly a non-renormalization theorem [13] is proven. According to this theorem the coefficient we computed does not receive corrections to higher orders approximation. To be specific the statement is that a renormalization scheme of the composite operators does exist such that (18.10) is exact, and it has been checked at two and three loop approximation, in abelian and for the singlet current in non-abelian theory [14–17]. The sum of the two loop triangular graphs turns out to be finite, exactly as in our computation, but in order to verify the non-renormalization theorem the one-loop counterterms have to be computed using the definition of γ_5 of (18.23), and imposing the validity of (18.10) in the $\psi - \bar{\psi}$ sector.

In our discussion the inconsistency of NDR was related to the anomaly computation. There are however simple formal arguments to prove it. Let's assume for a moment the anticommutation rule $\{\gamma^\mu, \gamma_5\} = 0$ in DR. Then we have

$$D\text{tr}[\gamma_5] = \text{tr}[\gamma_5\gamma^\mu\gamma_\mu] = \text{tr}[\gamma_\mu\gamma_5\gamma^\mu] = -\text{tr}[\gamma_5\gamma_\mu\gamma^\mu] = -D\text{tr}[\gamma_5]$$

Having “in some sense” the trace an analytic dependence on D one would conclude $\text{tr}[\gamma_5] = 0$.

In similar way

$$\begin{aligned} D\text{tr}[\gamma_5\gamma^\rho\gamma^\sigma] &= \text{tr}[\gamma_5\gamma^\rho\gamma^\sigma\gamma^\mu\gamma_\mu] = \text{tr}[\gamma_\mu\gamma_5\gamma^\rho\gamma^\sigma\gamma^\mu] = \\ &= -\text{tr}[\gamma_5\gamma_\mu\gamma^\rho\gamma^\sigma\gamma^\mu] = -2\text{tr}[\gamma_5\{\gamma^\rho, \gamma^\sigma\}] + (4 - D)\text{tr}[\gamma_5\gamma^\rho\gamma^\sigma] = \\ &= (4 - D)\text{tr}[\gamma_5\gamma^\rho\gamma^\sigma]. \end{aligned}$$

Where we used also (18.5) and the vanishing of $\text{tr}[\gamma_5]$. Then we arrive at $(2 - D)\text{tr}[\gamma_5\gamma^\rho\gamma^\sigma] = 0$, and the same analytic argumentation leads to $\text{tr}[\gamma_5\gamma^\rho\gamma^\sigma] = 0$.

Now the same technique yields $(4 - D)\text{tr}[\gamma_5\gamma^\rho\gamma^\sigma\gamma^\mu\gamma^\nu] = 0$. The content of this equation is that in $D = 4$ the trace can be different from zero, as it is, but in DR we perform limits as $D \rightarrow 4$, and then $D \neq 4$, then again $\text{tr}[\gamma_5\gamma^\rho\gamma^\sigma\gamma^\mu\gamma^\nu] = 0$. The conclusion should be that if the trace $\text{tr}[\gamma_5\gamma^\rho\gamma^\sigma\gamma^\mu\gamma^\nu]$ arises in Feynman graph computations in $D = 4$, then NDR cannot be used consistently in DR.

18.6 Conclusion

We have considered the problem of the definition of γ_5 in DR. It turns out to be strongly restricted by the axial anomaly. Indeed the axial anomaly coefficient can be computed in DR without explicit reference to any algebra of γ_5 , but only by using the fact that the traces assume their usual values in $D = 4$.

Extension of the D -dimensional algebra to γ_5 seems in any case necessary for the two-loop check of the non-renormalization theorem of the anomaly.

We have considered the 't Hooft and Veltmann solution to this problem. Being proved its mathematical and physical consistency it is widely accepted since time.

This could be the end of the story, but the subject is more involved. Equations (18.8) and (18.9) describe a global symmetry, but chiral currents may couple to gauge bosons. This happens when chiral fermions transform under a gauge group, that is the case of the Standard Model of Interactions.

In these models the anomaly must cancel (for suitable transformation laws of the fermions) so that the theory does exist as a quantum field theory, and this is only a part of the Ward identities, which must be fulfilled.

The use of 't HV prescription for this task is possible [18] but unwieldy. Indeed not critical sectors of the Ward identities, which would be automatically satisfied with an anticommuting γ_5 (NDR), lead now to “spurious anomalies” which have to be subtracted by hands, for instance loops with an even number of γ_5 .

As a matter of fact, without rigorous proofs, a wisdom on use of dedicated recipes for the anomalous graphs and otherwise the NDR scheme, has developed since first years [19, 20], see [21] and references therein for a more modern and comprehensive review.

References

1. N.N. Bogoliubov, D.V. Shirkov, *Introduction to the Theory of Quantized Fields*, (Wiley-Interscience Publication, 1980)
2. G.'t Hooft, M. Veltman, Nucl. Phys. B **44**, 189 (1972)
3. G.C. Bollini, J.J. Giambiagi, Nuovo Cim. **12 B**, 20 (1972)
4. G.M. Cicuta, E. Montaldi, Nuovo Cim. Lett. **4**, 329 (1972)
5. D.A. Akyeampong, R. Delbourgo, Nuovo Cim. **17 A**, 578 (1973)
6. P. Breitenlohner, D. Mason, Comm. Math. Phys. **52**, 11 (1977)
7. P. Breitenlohner, D. Mason, Comm. Math. Phys. **52**, 39 (1977)
8. P. Breitenlohner, D. Mason, Comm. Math. Phys. **52**, 55 (1977)
9. S.L. Adler, Phys. Rev. **177**, 2426 (1969)
10. J.S. Bell, R Jackiw, Nuovo Cim. **LX A**, 47 (1969)
11. R. Ferrari, Managing γ_5 in dimensional regularization II: the trace with more γ_5 , hep-th/150307410. Int. J. Theor. Phys. **56**, 691 (2017)
12. C. Itzykson, J.B. Zuber, *Quantum Field Theory* (McGraw-Hill International, 1980)
13. S.L. Adler, W.A. Bardeen, Phys. Rev. **182**, 1517 (1969)
14. D.R.T. Jones, J.P. Leveille, Nucl. Phys. B **206**, 473 (1982)
15. S.A. Larin, The renormalization of the axial anomaly in dimensional regularization, hep-ph/9302240. Phys. Lett. B **303**, 113 (1993)
16. M. Bos, Explicit Calculation of the Renormalized Singlet Axial Anomaly hep-ph/9211319. Nucl. Phys. B **404**, 215 (1993)
17. F. Jegerlehner, O.V. Tarasov, Explicit results for the anomalous three point function and non-renormalization theorems hep-ph/0510308. Phys. Lett. B **639**, 299 (2006)
18. G. Bonneau, Phys. Lett. **96 B**, 147 (1980) . G. Bonneau, Nucl. Phys. B **177**, 523 (1981) . G. Bonneau, Int. J. Mod. Phys. A **5**, 3831 (1990)
19. W.A. Bardeen, R. Gastmans, B. Lautrup, Nucl. Phys. B **46**, 319 (1972)
20. M. Chanowitz, M. Furman, I. Hinchliffe, Nucl. Phys. B **159**, 225 (1979)
21. F. Jegerlehner, Facts of life with γ_5 hep-th/0005255. Eur. Phys. J. C **18**, 673 (2001)

Further Reading

22. J. Collins, *Renormalization*. (Cambridge Monographs on Mathematical Physics, 1984)

Chapter 19

The Neutrino: Looking Through Its Experimental World



Alessandra Carlotta Re

Abstract At present, neutrino are still one of the most elusive and mysterious particles belonging to the so-called Standard Model of particles Physics. In this contribution I will present three international experiment (Borexino, SOX and JUNO) that are already (or will soon) contributing in a better understanding of the neutrino's features. The Borexino experiment is located in Hall C of the Laboratori Nazionali del Gran Sasso (Italy) and its main goal is the study of the properties of low energy solar neutrinos. The SOX experiment aims at the complete confirmation or at a clear disproof of the existence of sterile neutrino components. If successful, SOX will demonstrate the existence of at least one sterile neutrino component and will open a brand new era in fundamental particle physics and cosmology. The JUNO experiment, at present under-construction in Kaiping, China, has been designed in order to succeed in determining the neutrino mass hierarchy as well as to precisely measure the neutrino oscillation parameters.

19.1 Introduction

Neutrino's fame has incredibly grown in recent years. This is mainly due to the exciting developments that neutrino physics has seen over the past two decades, following the experimental evidence of their oscillations in the atmospheric, solar, reactor and accelerator sectors. In fact, a precise measurement of neutrino properties plays a crucial role in understanding many phenomena, from fundamental interactions to physics at the cosmological scale.

The basic nuclear process in the sun and in most of the other stars, is the fusion of hydrogen into helium: electron neutrinos are key-products of this nuclear reactions' chain. A precise measurement of solar neutrino fluxes yields on the one hand, if

A. C. Re (✉)

Dipartimento di Fisica "Aldo Pontremoli", Università degli Studi di Milano, Via Celoria 16, 20133 Milan, Italy

e-mail: alessandra.re@mi.infn.it

A. C. Re

INFN, Sezione di Milano, Via Celoria 16, 20133 Milan, Italy

© Springer Nature Switzerland AG 2018

P. F. Bortignon et al. (eds.), *Toward a Science Campus in Milan*,
https://doi.org/10.1007/978-3-030-01629-6_19

235

Standard Solar Model (SSM) predictions are assumed, to pin-down the neutrino survival probability at different energies (probing the MSW-LMA solution to the oscillation model and eventually setting constraints on possible NSI). On the other hand, if oscillation parameters are assumed, the study of specific solar neutrino components can cross-check the SSM predictions.

In this contribution three international neutrino physics experiments (Borexino, SOX and JUNO) and their physics goals are introduced. Existing and future experiments for detection of low-energy neutrinos or rare events need a low-background environment and thus are located in a deep underground site so to reduce the cosmic radiation background by several order of magnitude. This is the reason for which both Borexino-SOX and Juno are (will be) running in the underground laboratories of Gran Sasso (Italy) and Jiangmen (China) respectively.

19.2 The Borexino Experiment

The fusion processes that generate the energy of the Sun take place in the core, its innermost part. Lots of photons and neutrinos are produced there but while photons are stopped in the Sun core and the relative energy takes typically a million years to reach the Sun's surface, neutrinos emitted in the fusion reactions are able to instantaneously traverse the material of the Sun (mostly ionized hydrogen and helium) and arrive in 8 mins to the Earth's surface. This is possible since the neutrino cross-section is unbelievably small ($\approx 10^{-46}$ cm²): most of the neutrinos traverse the Sun and the Earth with almost no interactions at all.

It is precisely inside a huge detector like Borexino that an impressive compensation takes place between an unbelievably small number (that is the neutrino cross section) and an incredibly large number (the solar neutrino flux, $\approx 10^{10}$ cm⁻² s⁻¹) to produce a number which is still rather small ($\approx 10^{-36}$ interactions per second per target unit) but can be made reasonable by having a big (several tons of material) detector. Borexino's main goal has been to precisely measure the flux of ⁷Be low energy solar neutrinos. The Borexino detector features about 1200 tons of scintillator and 2800 tons of ultra-pure water as additional shielding against cosmic and natural radioactivity. A schematic view of the Borexino detector can be seen in Fig. 19.1.

In Borexino the detection of neutrinos is carried out by elastic scattering on the electrons of the liquid scintillator via the reaction: $\nu + e^- \rightarrow \nu + e^-$. Following the neutrino interaction, the target electron (previously practically at rest) received energy from the impinging neutrino and is therefore able to generate a scintillation light signal. This light burst is then detected by the detector's photomultipliers yielding the signature of a neutrino interaction.

The success of Borexino comes as a result of a 15 year long R&D study carried out by the collaboration to develop the best techniques of scintillator purification, allowing to reach and exceed the required levels of radiopurity. At present, Borexino is the less radioactive place in the world!

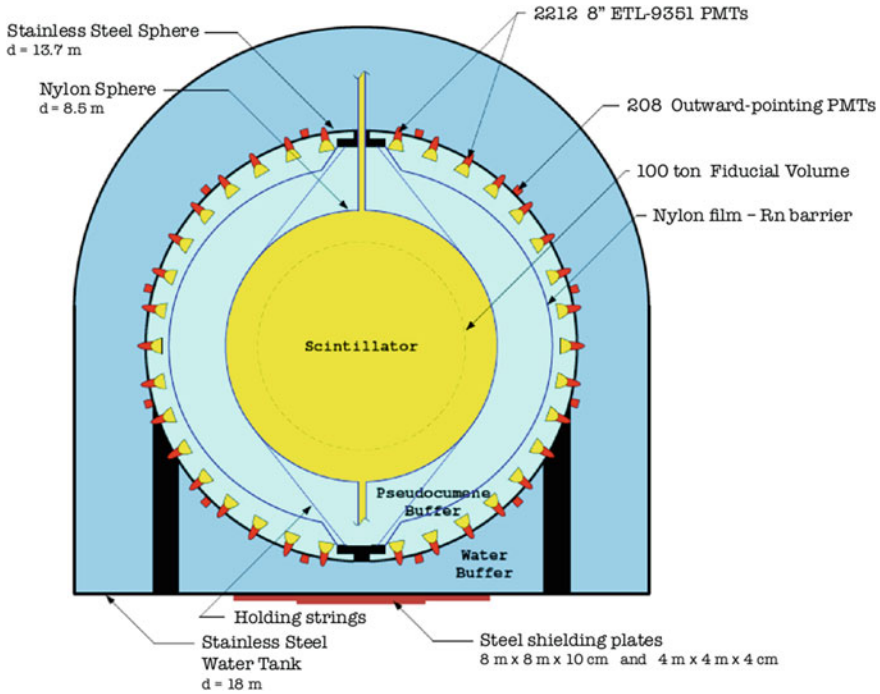


Fig. 19.1 Schematic view of the Borexino detector. An external water shielding surrounds the photomultiplier sphere which in turns contains about 1200 tons of an aromatic-based liquid scintillator

The Borexino experiment started taking data in 2007 and, since then, it has produced a considerable amount of interesting results. First of all it targeted its main goal: a precision measurement of the ^7Be solar neutrino rate [1] (with a total error of less than 5%) as well as a precise measurement of the ^7Be signal day-night asymmetry [2] were published in the early 2010s. Moreover, Borexino turned out to be an incredible multi-purpose experiment: in the last years its collaboration released the first direct spectroscopy of proton-proton solar neutrinos [3], the first direct measurement of the so-called pep solar neutrinos [4] and the measurement of the ^8B solar neutrino rate with an unprecedented low energy threshold [5]. Borexino has also published significant results on non-solar neutrino physics, such as the first observation of anti-neutrinos from the Earth (the geoneutrinos) and several limits on rare or forbidden processes.

Besides its application in the solar physics and geophysics fields, the Borexino detector offers a unique opportunity to perform a short-baseline neutrino oscillation study. This is the idea of the SOX (Short distance neutrino Oscillations with boreXino) project.

19.3 The SOX Experiment

The SOX experiment [6] aims at the complete confirmation or at a clear disproof of the so-called neutrino anomalies, a set of circumstantial evidences of electron neutrino disappearance observed with the LSND [7] and MiniBoone [8] experiments, with nuclear reactors [9, 10] and with solar neutrino Gallium detectors [11, 12]. If successful, SOX will demonstrate the existence of sterile neutrino components and will open a brand new era in fundamental particle physics and cosmology. A solid signal would mean the discovery of the first particles beyond the Standard Electroweak Model and would have profound implications in our understanding of the Universe and of fundamental particle physics. In case of a negative result, SOX would be able to close a long-standing debate about the reality of the neutrino anomalies, would probe the existence of new physics in low energy neutrino interactions and would provide a measurement of the neutrino magnetic moment.

The SOX experiment will use a powerful and innovative antineutrino generator made of $^{144}\text{Ce}-^{144}\text{Pr}$. That generator will be located at a short distance (about 8.5 m) from the center of the Borexino detector (in a pit beneath the Borexino detector tank) and will yield about ten thousands of clean antineutrino interactions in the internal volume of the Borexino detector. Measuring the interaction position (i.e. the source-detection point distance, L) and energy (E), an oscillated signature will be observed in case of a sterile neutrino according to the survival probability relation that, for sake of simplicity, can be expressed as:

$$P(\bar{\nu}_e \rightarrow \bar{\nu}_s) \propto \sin^2 \left(\frac{\pi}{2.48} \frac{L}{E} \Delta m_{41}^2 \right)$$

SOX will adopt two complementary and simultaneous strategies: a rate analysis and a shape analysis. The rate analysis will consist in counting and comparing the total rate of events with the expected one; it will be challenging since it critically depends on the accuracy of the expected rate estimation namely on the precise knowledge of the activity of the source, on the detection efficiency and on the determination of the active scintillator volume. A second analysis will be carried out by studying the oscillometry of the anti-neutrino signals. It will scan the events pattern as a function of the source-detection distance (L) and energy (E) looking for any shape deformation due to $\bar{\nu}_e - \bar{\nu}_s$ oscillations. This special analysis can be performed thanks to the Borexino high precision in position and energy reconstruction ($\sigma_E/E \approx 10\%$, $\sigma_L \approx 10 \text{ cm} @ 1\text{MeV}$). Figure 19.2 shows the expected oscillation pattern in absence (panel left) or presence (panel right) of electron to sterile neutrino oscillations, assuming as oscillation parameters $\Delta m_{41}^2 = 2.5 \text{ eV}^2$ and $\sin^2 2\theta = 0.14$.

The SOX experiment is starting in spring 2018 and will take data for about two years: the favourable combination of distance and energy ranges will allow SOX to probe a vast active-sterile Δm_{41}^2 range, spanning from 0.1 to 10 eV^2 .

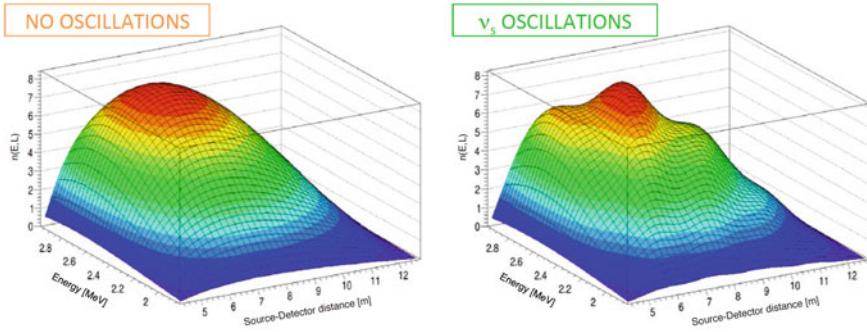


Fig. 19.2 The expected SOX oscillometry pattern as a function of the reconstructed energy and distance of the event from the source, assuming as oscillation parameters $\Delta m_{41}^2 = 2.5 \text{ eV}^2$ and $\sin^2 2\theta = 0.14$. The no oscillations case is depicted in panel left while the oscillated case is in panel right

19.4 The JUNO Experiment

The Jiangmen Underground Neutrino Observatory (JUNO, [13, 14]), a multipurpose liquid scintillator detector located in the underground facility of Kaiping (China), has been designed having as primary physics goal the determination of the neutrino mass hierarchy at high statistic significance ($3-4\sigma$). Nevertheless, the excellent energy resolution and a large fiducial volume seem to offer exciting opportunities for addressing many important topics in neutrino and astroparticle physics.

In addition to the neutrino mass hierarchy discrimination, JUNO's claimed goals are the precise determination of the neutrino oscillation parameters ($\sin^2 \theta_{12}$, Δm_{21}^2 , and $|\Delta m_{ee}^2|$ to an accuracy of better than 1%), the observation of antineutrinos/neutrinos from terrestrial and extra-terrestrial sources (i.e. geoneutrinos or supernova, solar, and atmospheric neutrinos) as well as the identification of sterile neutrinos (if any) having Δm_{41}^2 in the $[10^{-5}, 10^{-2}] \text{ eV}^2$ range and a sufficiently large mixing angle θ_{14} . This latter analysis will be carried out through a precise measurement of the reactor antineutrino energy spectrum (Fig. 19.3). JUNO will also be sensitive to several nucleon decays and will look for neutrinos resulting from dark-matter annihilation in the Sun.

JUNO will detect antineutrinos generated by a cluster of nuclear power plants ($\approx 17.5 \text{ GW}$ total power) approximately located at 50 km from the detector site. At present the detector is under construction and is expected to be filled and start taking data in early 2020.

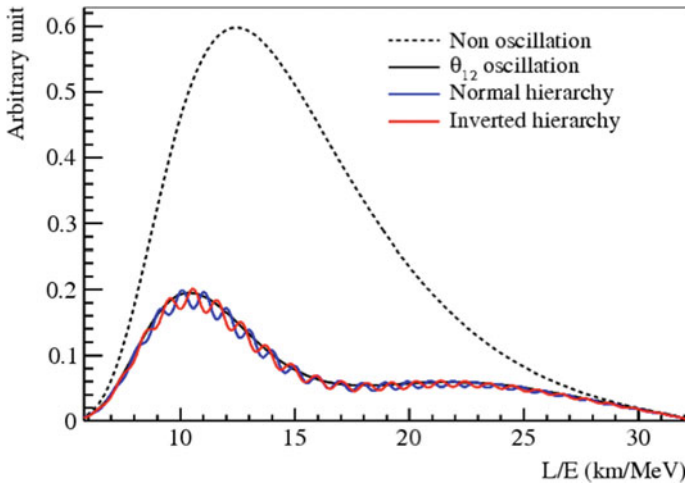


Fig. 19.3 The expected relative shape difference of the reactor antineutrino flux according to different neutrino mass hierarchy. The dashed line represents the no-oscillation case; the red and blue lines describe the oscillation pattern according to the inverted and normal mass hierarchy respectively

19.5 Conclusions

Borexino has been, so far, the only experiment able to perform the real-time spectroscopy of neutrinos from almost all the nuclear reactions happening in the sun. The goals of the second phase of the Borexino solar neutrino program are mainly focused around improving its current limit of the CNO neutrino flux and possibly measure it.

In parallel to this program, the Borexino detector will be the base for the SOX project, a short baseline experiment, aiming at definitively investigating the sterile neutrino hypothesis. The SOX experiment is starting in spring 2018.

Lastly, in near future, the neutrino community will have exciting news from the JUNO experiment whose impressive scientific program spans from the determination of the neutrino mass hierarchy to the precise measurement of the neutrino oscillation parameters, and to dark-matter search.

References

1. G. Bellini et al., (Borexino collaboration): Precision measurement of the ^7Be solar neutrino interaction rate in Borexino. *Phys. Rev. Lett.* **107**, 141302 (2011)
2. G. Bellini et al., (Borexino collaboration): Absence of a day-night asymmetry in the ^7Be solar neutrino rate in Borexino. *Phys. Lett. B* **707**, 22 (2012)
3. G. Bellini et al., (Borexino collaboration): Neutrinos from the primary proton-proton fusion process in the Sun. *Nature* **512**, 383–386 (2014)

4. G. Bellini et al., (Borexino collaboration): First evidence of pep solar neutrinos by direct detection in Borexino. *Phys. Rev. Lett.* **108**, 051302 (2012)
5. G. Bellini et al., (Borexino collaboration): Measurement of the solar ^8B neutrino rate with a liquid scintillator target and 3 MeV energy threshold in the Borexino detector. *Phys. Rev. D* **82**, 033006 (2010)
6. G. Bellini et al., (Borexino collaboration): SOX: short distance neutrino oscillations with BoreXino. *JHEP* **8**, 038 (2013)
7. A. Aguilar et al., (LSND Collaboration): Evidence for neutrino oscillations from the observation of electron anti-neutrinos in a muon anti-neutrino beam. *Phys. Rev. D* **64**, 112007 (2001)
8. A. Aguilar et al., (MiniBooNe Collaboration): improved search for $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ oscillations in the MiniBooNE experiment. *Phys. Rev. Lett.* **110**, 161801 (2013)
9. G. Mention et al., Reactor antineutrino anomaly. *Phys. Rev. D* **83**, 073006 (2011)
10. F.P. An et al., (Daya Bay Collaboration): Evolution of the reactor antineutrino flux and spectrum at Daya Bay. *Phys. Rev. Lett.* **118**, 25180 (2017)
11. W. Hamper et al., (Gallex Collaboration): Final results of the ^{51}Cr neutrino source experiments in GALLEX. *Phys. Lett. B* **420**, 114–126 (1998)
12. J.N. Abdurashitov et al., (SAGE Collaboration): Results from SAGE (The Russian-American gallium solar neutrino experiment). *Phys. Lett. B* **328**, 234–248 (1994)
13. F.P. An et al., (Juno Collaboration): Neutrino physics with JUNO. *J. Phys. G* **43**, 030401 (2016)
14. JUNO, The Jiangmen Underground Neutrino Observatory, <http://juno.ihep.cas.cn>

Chapter 20

Towards a Universal Nuclear Structure Model



X. Roca-Maza, Pier Francesco Bortignon, G. Colò, Y. F. Niu and E. Vigezzi

Abstract We present part of our recent developments of a microscopic effective model for a consistent description of both bulk and single-particle properties of nuclei.

20.1 Introduction

A quantitative theoretical description of atomic nuclei continues to represent a major challenge, in spite of the intense activity of the nuclear physics community over more than 70 years [1]. On the one hand, the complexity of the problem is related to the difficulty to derive the nucleon-nucleon (NN) interaction by first principles from QCD, which is non-perturbative at low-energies of interest (from hundreds of keV to tens of MeV) [2]; on the other hand, for a given interaction, a complete solution of a strongly interacting system of self-bound neutrons and protons is not yet available [3].

The use of NN-potentials fitted to reproduce available NN-scattering data in the vacuum constitutes one of the current theoretical possibilities, which however

Pier Francesco Bortignon has deceased.

X. Roca-Maza (✉) · P. F. Bortignon · G. Colò · Y. F. Niu
Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano,
Via Celoria 16, 20133 Milan, Italy
e-mail: xavier.roca.maza@mi.infn.it
URL: <http://www.mi.infn.it/~roca/>

X. Roca-Maza · P. F. Bortignon · G. Colò · E. Vigezzi
INFN, Sezione di Milano, Via Celoria 16, 20133 Milan, Italy

Y. F. Niu
ELI-NP, Horia Hulubei National Institute for Physics and Nuclear Engineering,
30 Reactorului Street, 077125 Bucharest-Magurele, Romania

Y. F. Niu
School of Nuclear Science and Technology, Lanzhou University,
730000 Lanzhou, China

presents a limited applicability to systematic studies of nuclei along the whole nuclear chart (cf. Table I of [4]).

The Milano group is instead focused on approaches based on the use of phenomenological effective potentials fitted to experimental data on selected nuclei. Once the parameters of such interaction have been determined, the same interaction can be used to make predictions along the mass table. Such models constitute a unique tool to theoretically access some ground and excited state properties for all of the nuclei that have been measured in the past and that will be measured in the next future in Rare Ion Beam Facilities world wide [5].

These effective interactions are sufficiently soft to allow the use of perturbation theory. At first order (Hartree-Fock), one can determine a self-consistent static mean field. The same effective interaction can also be employed to calculate the properties of collective excitations, usually within the Random Phase Approximation (RPA). The Milano group has worked both on the development of these interactions, devoting particular attention to spin and isospin channels [6], and on the accurate self-consistent solution of the HF+RPA problem [7].

These models can also be understood as approximate realizations of an exact nuclear energy density functional (EDF). By construction, nuclear EDF approaches are only reliable for the description of nuclear masses and densities or, more in general, for the expectation value of one-body operators, while the associated single-particle dynamics is not realistic. To make further progress, one needs to go to higher order in perturbation theory. However, to blindly apply such a strategy is very complicated from the technical point of view and may not entail, in general, a clear physical interpretation of the considered diagrams. Hence, it can be a crucial advantage to individuate the relevant degrees of freedom of the problem under consideration and to connect them with given diagrams, or series of diagrams, in the many-body expansion. A possible scheme is offered by Nuclear Field Theory and similar approaches, which provide a consistent perturbation scheme based on the interplay of single-particle degrees of freedom with collective states. When such particle-vibration coupling (PVC) is taken into account, one can obtain already at a low order in the perturbative expansion a quantitative description of the fragmentation of single-particle states, as probed by one-nucleon transfer reactions, and of the observed strength functions associated with giant collective modes.

20.2 Theoretical Framework

In this section, we will briefly describe the bases of our formalism paying attention to the underlying physical assumptions and refer the reader to the references herein for more technical details.

In general, the nuclear Hamiltonian can be written as $\mathcal{H} = T + V$ where T represents the kinetic energy and V the two-body¹ effective interaction between nucleons. Adding and subtracting an auxiliary one-body potential U to \mathcal{H} allows us to formulate the problem in terms of a non-interacting part $\mathcal{H}_0 \equiv T + U$ that corresponds to the so called Hartree-Fock (HF) Hamiltonian if the auxiliary potential is defined as the ground state expectation value of V on a Slater determinant, plus a term $V - U$ that vanishes by construction within the HF approximation. That is, the solution of \mathcal{H}_0 coincides with that of \mathcal{H} in first order perturbation theory. It can be shown that this approximation allows to write the total energy of the system in terms of the relevant densities (specifically: nucleon densities, spin densities and currents) and, therefore, allows for its interpretation as an approximate realization of an EDF. The Milano group has been very active in developing effective interactions suitable for the self-consistent calculations at the mean field level of bulk nuclear properties and of giant collective modes [8–10]. The latter can be calculated as superpositions of 1-particle 1-hole (1p-1h) excitations by solving the linear response equations for small perturbations obtained from the time-dependent Schroedinger equation. Such an approach is known as the Random Phase Approximation (RPA), and is one of the most common ways to build the nuclear vibrations or phonons in a fully consistent way based on the Hartree-Fock ground state.

In order to calculate in detail the shape of the strength functions associated with giant resonances, one needs to go beyond the RPA. In spherical nuclei, a good fraction of the observed width can be obtained taking into account the coupling of collective low-energy nuclear vibrations to single-particle degrees of freedom.

It is customary to work within a subspace \mathcal{Q}_1 that contains all nuclear 1p-1h configurations built with the single-particle states, eigenfunctions of the HF Hamiltonian, where the RPA will correspond to solve the initial Hamiltonian projected in the \mathcal{Q}_1 subspace $\mathcal{Q}_1 \mathcal{H} \mathcal{Q}_1$. The RPA states built in this way are linear combinations of 1p-1h (HF) states that represent, as mentioned, a nuclear vibration or phonon. In general, the coupling of nuclear vibrational states with single-particle states (PVC) leads to the transfer of energy from one degree of freedom to another. This allows for the rearrangement of the internal degrees of freedom giving rise to a damping via the so called spreading width (Γ^\downarrow). A PVC approach provides also a more realistic description—when compared to the HF or RPA results—of the emission of a γ ray and the escape of nucleons which contributes to the so called escape width (Γ^\uparrow).

As in [11], in order to model the escape and spreading widths, we define two additional sub-spaces. The first one \mathcal{P} is made of holes plus unbound HF states which have positive energy and that we construct to be orthogonal to the bound occupied and unoccupied HF states. The second subspace \mathcal{Q}_2 is made of 1p-1h excitations $|ph\rangle$ coupled to a phonon state labeled as $|N\rangle$. Note that by construction

¹Three-body forces have been shown to be necessary as well for a good description of nuclear phenomenology such as the saturation density and energy. Within approaches based on density functional theory, it is customary to approximate a three body force by a two-body density dependent force.

$\mathcal{Q}_1 \mathcal{Q}_2 = 0$ since $\langle ph|N \otimes ph \rangle$ will be always zero. Finally, we define the subspace $\mathcal{Q} = \mathcal{Q}_1 + \mathcal{Q}_2$.²

Using the properties of the sub-spaces defined here and following a method very similar to that of [12], one can write \mathcal{H} projected in the \mathcal{Q}_1 subspace as

$$\begin{aligned} \mathcal{H}_{\mathcal{Q}_1} &= \mathcal{Q}_1 \mathcal{H} \mathcal{Q}_1 \\ &+ \mathcal{Q}_1 \mathcal{H} \mathcal{P} \frac{1}{\mathcal{P}(\omega - \mathcal{H} - i\epsilon) \mathcal{P}} \mathcal{P} \mathcal{H} \mathcal{Q}_1 \\ &+ \mathcal{Q}_1 \mathcal{H} \mathcal{Q}_2 \frac{1}{\mathcal{Q}_2(\omega - \mathcal{H} - i\epsilon) \mathcal{Q}_2} \mathcal{Q}_2 \mathcal{H} \mathcal{Q}_1 \\ &+ \text{higher order terms.} \end{aligned} \quad (20.1)$$

where the first term in the right hand side of the equation corresponds to the RPA solution, $\mathcal{Q}_1 \mathcal{H} \mathcal{Q}_1$; the second term corresponds to excite a bound particle or hole state to the continuum, its propagation on top of the non-interacting (HF) potential for states with positive energy and its deexcitation to a bound state again. This term, commonly labeled as \mathcal{W}^\dagger , will produce the escape width previously discussed. The third term corresponds to the coupling of a particle or hole state with more complex configurations represented by \mathcal{Q}_2 , its propagation on top of the potential, and the reabsorption of this complex state into a particle or a hole. This term, commonly labeled as \mathcal{W}^\downarrow , will produce the spreading width previously discussed. Higher order terms are neglected in our calculations due to the reasons discussed in [13].

In what follows, we define some useful and related quantities. The observed spectrum of a nucleus excited by an external field \mathcal{F} is described by the nuclear polarization propagator or dynamic polarizability $\Pi(\omega)$. It corresponds to the double convolution with \mathcal{F} of the propagator or response function, that is

$$\Pi(\omega) = \langle 0 | \mathcal{F}^\dagger \frac{1}{\omega - \mathcal{H}_{\mathcal{Q}_1}(\omega) + i\epsilon} \mathcal{F} | 0 \rangle. \quad (20.2)$$

The strength function is defined as

$$\mathcal{S}(\omega) = -\frac{1}{\pi} \Im [\Pi(\omega)]. \quad (20.3)$$

20.3 Results

In this section we will present some selected results, concerning first calculations at the mean field level, and then calculations performed within the particle-vibration scheme.

²The projectors \mathcal{Q} and \mathcal{P} cover the full model space: $\mathcal{Q} + \mathcal{P} = 1$ and $\mathcal{Q}\mathcal{P} = 0$; and all the projectors fulfill the usual conditions: $\mathcal{P}^2 = \mathcal{P}$, $\mathcal{Q}^2 = \mathcal{Q}$ and $\mathcal{Q}_i^2 = \mathcal{Q}_i$.

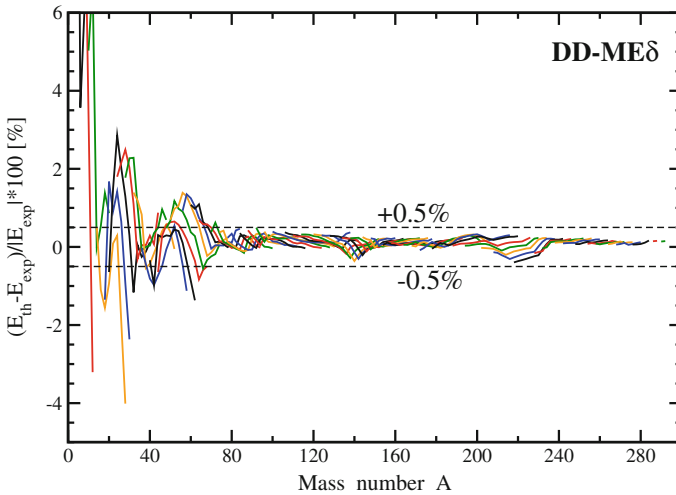


Fig. 20.1 Relative difference in % between the experimental and theoretical binding energies as predicted by DD-ME δ . Figure taken from [14]

Mean-field approach. Some of us have built two new nuclear structure models able to compete with currently available models and to improve on them in certain aspects. The first one is a relativistic model named DD-ME δ [15]. It is based on a Lagrangian where effective mesons carry the interaction. The novelty of this model is that it reproduces state-of-the-art ab initio calculations for the nuclear matter equation of state [16, 17]—energy against density of an infinite system of neutrons and protons—and at the same time provides a satisfactory description of the ground state properties in finite nuclei such as binding energies (cf. Fig. 20.1). As it can be seen in Fig. 20.1, the accuracy of this model is better than 0.5% for nuclei heavier than $A = 80$, where A stands for the mass (baryon) number, while it is less accurate for light nuclei. This is at the level of accuracy of other models of the same type available in the literature. In addition, it is clear from the arch structure of the residuals depicted in Fig. 20.1 that these types of models do not contain all the relevant physics to describe nuclear masses. Such arches appear also in other EDFs and are due to an incomplete description of the shell structure in nuclei. Hence, more sophisticated EDFs and/or many-body approaches will be needed to address a better description of this observable.

In a second model, some of the authors concentrated on improving the spin-isospin channel of the nuclear effective interaction while keeping all important properties at the same level of accuracy of commonly used nuclear mean-field models. We adopted a non-relativistic Skyrme ansatz for the nuclear effective interaction. The model was named SAMi [6] and it is able to improve the description of spin-isospin resonances [18] when compared to other EDFs. As an example, we show in Fig. 20.2 the strength function of a spin-isospin resonance, known as Gamow-Teller resonance, in ^{208}Pb that in general plays an important role not only in nuclear physics studies but also in

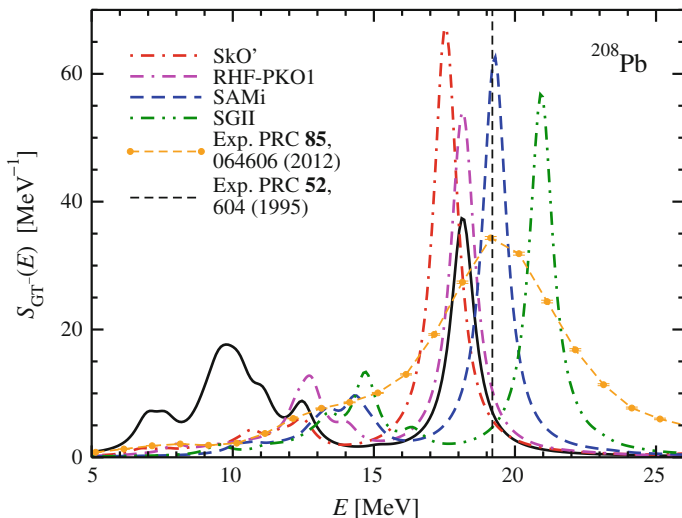


Fig. 20.2 Spin-isospin resonance (Gamow-Teller resonance) in ^{208}Pb as a function of the excitation energy. For details on the other models shown in this figure see [6]. Figure adapted from [6]

astrophysical processes [19, 20]. In this figure, SAMi is compared with other model predictions and experimental data and, as it can be seen, it perfectly reproduces the excitation energy of this resonance. For details on the theoretical or experimental data, we refer the reader to [6] and references therein.

PVC approach. In the most recent years we have improved the microscopic implementation of the PVC approach based on effective interactions of the Skyrme type. Some of the most relevant results are presented in what follows. For a more exhaustive overview we refer the reader to references herein.

In Fig. 20.3 we show the results for the β -decay half-life of the exotic ^{78}Ni which is an important waiting-point nucleus in the astrophysical r -process [24–26]. Most Skyrme interactions do not reproduce well this property (SAMi included). The inclusion of PVC effects reduces the half-lives systematically improving the overall agreement between the different models and experiment.

Next, we show in Fig. 20.4 the effects of the implementation of the so called *subtraction method* [27] that can be regarded as a recipe to renormalize the interaction at the particle-vibration coupling vertex. The renormalization of our model is needed for two reasons [28, 29]. On the one hand, the Skyrme interaction is zero-range. This implies that beyond mean field calculations—such as the ones based on the PVC approach—will show ultraviolet divergences, for example, in evaluating the matrix elements for the particle-vibration coupling vertex. On the other hand, it is customary to implement Skyrme interactions which have been fitted on experimental bulk properties at the mean field level, so that correlations explicitly included in the PVC approach had already been effectively included in the parameters of the interaction. This will introduce some double counting (see [29] for more details). In Fig. 20.4, we show the correction introduced by such method on the strength function

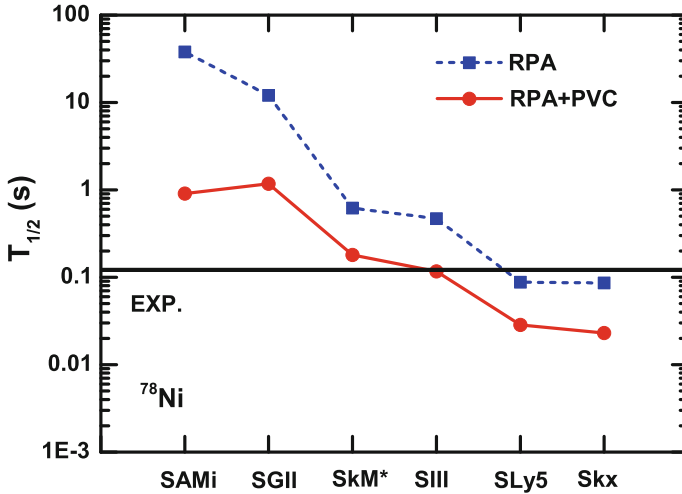


Fig. 20.3 β -decay half-life of ^{78}Ni , calculated by RPA and RPA+PVC approaches with different Skyrme interactions. Results from [21] with updated experimental data [22] are shown. Figure taken from [23]

for the giant quadrupole resonance in ^{208}Pb . As expected, the positions of the peaks including the *subtraction method* tend to be in much better agreement with the RPA results and experiment (black arrows in Fig. 20.3) than neglecting it [13, 27, 30].

Finally, in [31] pairing correlations were implemented in the PVC model in both the ground state and the excited states, making it possible to study the properties of resonances in open-shell nuclei. In Fig. 20.5 we show a calculation of the strength function of the Gamow-Teller resonance in ^{120}Sn calculated with the Quasiparticle RPA (QRPA) and Quasiparticle PVC (QPVC) models, without and with subtraction method, using the Skyrme interaction SkM*. As it has been seen in Fig. 20.4, the subtraction method shifts the strength function to larger excitation energies, closer to the QRPA results.

20.4 Conclusions

We have briefly presented some recent developments aimed at building a microscopic model for the description of both bulk and single-particle properties in nuclei. Models based on effective interactions solved at Hartree-Fock level are successful in the description of bulk properties in nuclei including not only the ground state but also some properties of collective excitations, in particular the energy of giant resonances. For a more accurate description of the fragmentation of single-particle states as well as of the width of giant resonances and other collective features one needs to go beyond the HF approach. The particle vibration coupling model is one

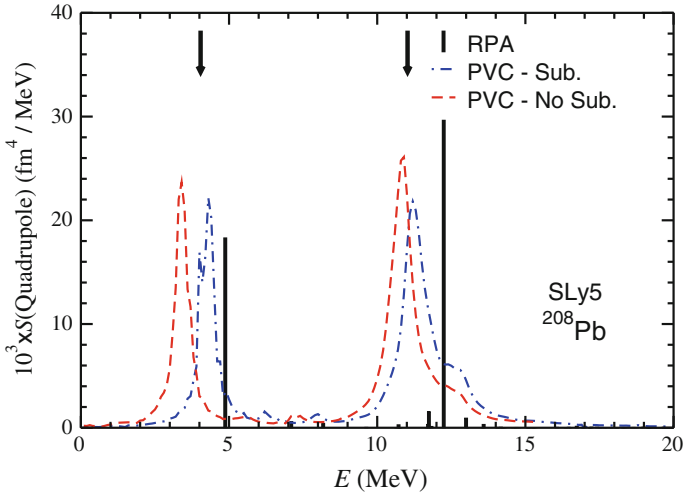


Fig. 20.4 The isoscalar quadrupole response in ^{208}Pb as predicted by SLy5. We show the RPA response in black bars, the PVC without subtraction in dashed red lines and the PVC with subtraction in dotted-dashed blue lines. One black arrow indicates the position of the measured 2_1^+ state, and the other black arrow indicates the experimental centroid energy $m_1/m_0 = 11.0 \pm 0.2$ MeV. Figure taken from [13]

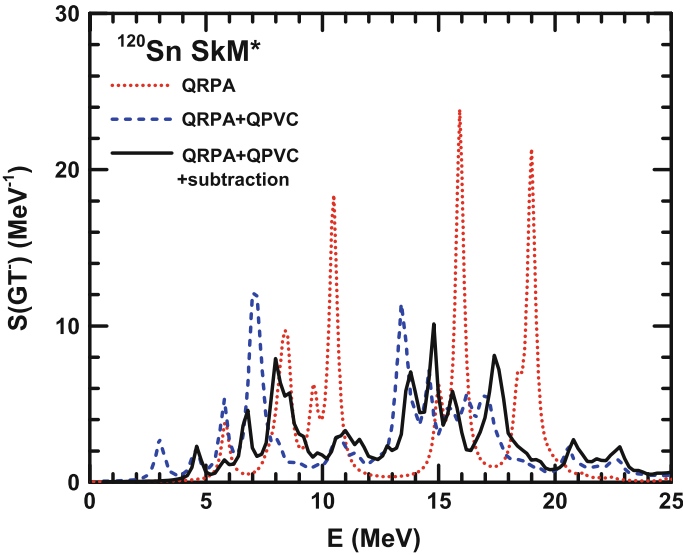


Fig. 20.5 The Gamow-Teller strength distributions for ^{120}Sn as predicted by Quasiparticle RPA (QRPA) and Quasiparticle PVC (QPVC) models, without and with subtraction method, using the Skyrme interaction SkM*. Figure taken from [31]

of the possible implementations of a beyond mean field theory that improves the description of these observables. Nevertheless, the renormalization of the interaction needs to be investigated. We have shown in [28, 29] for two simplified cases that the renormalization is possible. In addition, we have checked that the renormalization via the subtraction method for the full PVC model seems incomplete.

Acknowledgements This work was partly supported by Funding from the European Union’s Horizon 2020 research and innovation programme under Grant agreement No. 654002.

References










1. Special issue dedicated to the memory of Gerald E Brown (1926–2013). Nucl. Phys. A **928**, 9 (2014)
2. S. Aoki, T. Doi, T. Hatsuda, Y. Ikeda, T. Inoue, N. Ishii, K. Murano, H. Nemura, K. Sasaki, Prog. Theor. Exp. Phys. **2012**(1), 01A105 (2012)
3. P. Ring, P. Schuck, *The Nuclear Many-Body Problem* (Springer, 2004)
4. B.S. Hu, F.R. Xu, Q. Wu, Y.Z. Ma, Z.H. Sun, Phys. Rev. C **95**, 034321 (2017)
5. Y. Blumenfeld, T. Nilsson, P.V. Duppen, Phys. Scr. **2013**(T152), 014023 (2013)
6. X. Roca-Maza, G. Colò, H. Sagawa, Phys. Rev. C **86**, 031306 (2012)
7. G. Colò, L. Cao, N.V. Giai, L. Capelli, Comput. Phys. Commun. **184**(1), 142 (2013)
8. A. Bohr, B. Mottelson, *Nuclear Structure*. No. Vol. I & II in Nuclear Structure (World Scientific, 1998)
9. P.F. Bortignon, A. Bracco, R.A. Broglia, *Giant Resonances*. Contemporary concepts in physics (Taylor & Francis, 1998)
10. M. Harakeh, A. Woude, *Giant Resonances: Fundamental High-frequency Modes of Nuclear Excitation*. (Oxford Science Publications, Oxford University Press, 2001)
11. G. Colò, N. Van Giai, P.F. Bortignon, R.A. Broglia, Phys. Rev. C **50**, 1496 (1994)
12. S. Yoshida, Progress Theoret. Phys. Suppl. **74–75**, 142–156 1 Jan (1983). <https://doi.org/10.1143/PTPS.74.142>
13. X. Roca-Maza, Y.F. Niu, G. Colò, P.F. Bortignon, J. Phys. G: Nucl. Part. Phys. **44**(4), 044001
14. S.E. Agbemava, A.V. Afanasjev, D. Ray, P. Ring, Phys. Rev. C **89**, 054320 (2014)
15. X. Roca-Maza, X. Viñas, M. Centelles, P. Ring, P. Schuck, Phys. Rev. C **84**, 054309 (2011)
16. M. Baldo, C. Maieron, P. Schuck, X. Vias, Nucl. Phys. A **736**(3), 241 (2004)
17. M. Baldo, *Nuclear Methods and the Nuclear Equation of State* (World Scientific, 2011)
18. F. Osterfeld, Rev. Mod. Phys. **64**, 491 (1992)
19. H.A. Bethe, Rev. Mod. Phys. **62**, 801 (1990)
20. K. Langanke, G. Martínez-Pinedo, B. Müller, H.T. Janka, A. Marek, W.R. Hix, A. Juodagalvis, J.M. Sampaio, Phys. Rev. Lett. **100**, 011101 (2008)
21. Y.F. Niu, Z.M. Niu, G. Colò, E. Vigezzi, Phys. Rev. Lett. **114**, 142501 (2015)
22. Z.Y. Xu, S. Nishimura, G. Lorusso, F. Browne, P. Doornenbal, G. Gey, H.S. Jung, Z. Li, M. Niikura, P.A. Söderström, T. Sumikama, J. Taprogge, Z. Vajta, H. Watanabe, J. Wu, A. Yagi, K. Yoshinaga, H. Baba, S. Franchoo, T. Isobe, P.R. John, I. Kojouharov, S. Kubono, N. Kurz, I. Matea, K. Matsui, D. Mengoni, P. Morfouace, D.R. Napoli, F. Naqvi, H. Nishibata, A. Odahara, E. Şahin, H. Sakurai, H. Sakurafner, I.G. Stefan, D. Suzuki, R. Taniuchi, V. Werner, Phys. Rev. Lett. **113**, 032505 (2014)
23. Y. Niu, Z. Niu, G. Colò, E. Vigezzi, AIP Conf. Proc. **1681**(1), 050010 (2015)
24. E.M. Burbidge, G.R. Burbidge, W.A. Fowler, F. Hoyle, Rev. Mod. Phys. **29**, 547 (1957)
25. K. Langanke, G. Martínez-Pinedo, Rev. Mod. Phys. **75**, 819 (2003)
26. The Hans Bethe Centennial Volume 1906–2006. Phys. Rep. **442**(1), 237 (2007)
27. V.I. Tselyaev, Phys. Rev. C **88**, 054301 (2013)

28. M. Brenna, G. Colò, X. Roca-Maza, Phys. Rev. C **90**, 044316 (2014)
29. C.J. Yang, M. Grasso, X. Roca-Maza, G. Colò, K. Moghrabi, Phys. Rev. C **94**, 034311 (2016)
30. D. Gambacurta, M. Grasso, J. Engel, Phys. Rev. C **92**, 034303 (2015)
31. Y.F. Niu, G. Colò, E. Vigezzi, C.L. Bai, H. Sagawa, Phys. Rev. C **94**, 064328 (2016)

Chapter 21

Recent Advances in Scintillating Optical Fibre Dosimeters



Ivan Veronese , Marie Claire Cantone, Salvatore Gallo ,
Cristina De Mattia, Eduardo d'Ippolito, Nicola Ludwig ,
Marco Gargano , Simone Cialdi , Stefano Latorre, Norberto Chiodini ,
Mauro Fasoli , Federico Moretti , Eleonora Mones, Gianfranco Loi
and Anna Vedda 

Abstract Scintillating optical fibres have shown interesting results for ionizing radiation monitoring. Since they may enable a remote, punctual and real-time dose assessment, their application in medical dosimetry is very promising. This work aims to summarize some recent progresses in the development and characterization of rare-earth doped silica optical fibres. The radioluminescent and dosimetric properties of Ce, Eu and Yb-doped fibres are presented and the advantages and challenges in the use of these sensors for radiation therapy dosimetry are discussed. For such application, an effective approach to deal with the stem effect, i.e. the spurious luminescent signal originated in the light guide as a consequence of its exposition to ionizing radiations (i.e. Cerenkov light and intrinsic fluorescence) must be considered. The stem effect mainly occurs in the UV-VIS region. We demonstrated that the use of a dopant emitting in the near infrared, like Yb, is suitable for an optical discrimination of the dosimetric signal. Indeed, through a characterization of the dosimetric properties of

I. Veronese (✉) · M. C. Cantone · S. Gallo · C. De Mattia · E. d'Ippolito · N. Ludwig · M. Gargano
S. Cialdi
Dipartimento di Fisica "Aldo Pontremoli", Università degli Studi di Milano, Milan, Italy
e-mail: ivan.veronese@unimi.it

I. Veronese · M. C. Cantone · S. Gallo · S. Cialdi · S. Latorre
Istituto Nazionale di Fisica Nucleare, Sezione di Milano, Milan, Italy

N. Chiodini · M. Fasoli · F. Moretti · A. Vedda
Dipartimento di Scienza dei Materiali, Università di Milano-Bicocca, Milan, Italy

E. Mones · G. Loi
Medical Physics Department, Azienda Ospedaliera Maggiore della Carità, Novara, Italy

Present Address

M. C. Cantone
Dipartimento di Scienze Biomediche, Chirurgiche ed Odontoiatriche, Università degli Studi di
Milano, Milan, Italy

Present Address

F. Moretti
Lawrence Berkeley National Laboratory, Berkeley, USA

Yb-doped fibres in conjunction with an optical filter and an avalanche detector, we proved that the drawback due to the stem effect does not impair the system response even in the most challenging irradiation geometries, attesting to the robustness of the device in complex dosimetric scenarios.

21.1 Introduction

In the last decade, impressive technological improvements significantly changed the external radiation therapy treatments to a highest level of sophistication and complexity.

All these improvements could be a benefit for the patient, but at the same time open new challenges for the medical physicists, both in terms of strategies of treatment planning and in terms of radiation protection and dosimetry [1, 2]. Indeed, instruments and procedures that were suitable up to few years ago now may fail or provide results not sufficiently accurate.

In this contest, scintillating optical fibre dosimeters could be particularly useful in various applications [3]. Essentially, such dosimeter consists in a small scintillator that originates a radioluminescent (RL) signal when exposed to ionizing radiation. The scintillator is connected to an optical fibre acting as light guide to a suitable optical detector.

In principle, an optical fibre dosimeter should enable real time measurements of the dose rate in a specific point of the radiation field. The small dimensions could be exploited in small field dosimetry [4] and in vivo dosimetry. Moreover the fibre does not require high voltage and so it is safe its in vivo use on the patient body. Finally the radioluminescence is unaffected by magnetic field, and so fibre dosimeters could be particularly useful for quality assurance (QA) measurements in the new hybrid MRI-LINAC systems.

Actually, the interest in these systems dates back to the early '90 [5], but only almost 20 years later that the first commercial system became available, the Exradin W1 by Standard Imaging [6]. The main reason of this long history was the need to face efficiently the intrinsic drawback of fibre optics dosimeter, that is the so-called stem effect. The stem effect is an additional luminescent signal due to the irradiation of the optical fibre that contributes mainly in terms of Cerenkov light. It is a spurious signal since it strongly depends on the irradiation conditions, typically the length of fibre directly exposed to the radiation beam and the direction of incidence of the beam on the fibre axis.

Various methods for the correction or subtraction of the stem effect have been proposed in the literature, each presenting advantages but also limitations according to the irradiation scenario. A comprehensive discussion of the main approaches can be found elsewhere [3, 7, 8] and here briefly summarized. The simplest method consists in using a second (dummy) fibre for measuring the stem effect separately [5]. The peculiarities of the radiation beams used in the modern radiotherapy makes this approach not always reliable [9]. Another method consists in the temporal discrimina-

tion between the scintillation and the Cerenkov light [10]. A complete discrimination between the dosimetric signal (slower) and the spurious one (faster) can be therefore achieved, provided that the scintillating fibre is irradiated with a pulsed radiation beam. A third approach is based on the chromatic separation of the scintillation signal and the spurious one, by using proper optical filters and read out systems [11]. In such case, a complete optical separation cannot be achieved with conventional scintillators emitting in the visible region, therefore sophisticated calibration procedures are needed in order to get the suppression of the stem effect [12].

This work aims to summarize recent studies about the development and characterization of rare-earths doped silica optical fibre dosimeters, with particular emphasis to the research of new solutions for the stem effect drawback. These investigations were carried out in a frame of a collaboration between the Department of Physics of the University of Milan and the Department of Material Science of the University of Milano-Bicocca, exploiting clinical irradiation facilities and the expertise of various medical physicists, in primis, those of the University Hospital Maggiore della Carità of Novara.

21.2 Rare Earths Doped Silica Optical Fibres

The materials of interest were produced at the laboratories of the Department of Materials Science of the University of Milano-Bicocca, where a process to obtain rare-earths doped silica glassed, by sol-gel method, was developed. The procedure, described in detail in [13], is here briefly summarized.

Silica glasses with rare earth (RE) molar ppm concentration in the interval 0–5000 ppm were prepared by using Tetraethyl orthosilicate (TEOS, Aldrich, 99.999%) and $\text{RE}(\text{NO}_3)_3 \cdot 6\text{H}_2\text{O}$ (Aldrich, 99.99%) as precursors. The prepared SOL solution was sealed in polypropylene containers and stored in a thermostatic chamber at 35 °C for gelation. Such gelation process takes approximately 15 days. The resulting samples were aged for few days and, subsequently, small holes produced in the cover of the containers induced slow drying of the alcogel, yielding transparent xerogels. Densification of xerogels to glasses was obtained by a sintering procedure up to 1250 °C. Finally, a rapid thermal treatment (RTT) performed after densification proved to significantly improve the scintillation efficiency of the samples. RTT was obtained by using an oxidizing oxygen-hydrogen flame: after a very quick temperature increase (2–4 s), the sample was kept at approximately 1800 °C for 10 s and then rapidly cooled in air. An example of a scintillating doped silica glass sample, in form of disk is shown in Fig. 21.1a.

The described method for the production of bulk samples can be properly adjusted to obtain small glass cylinders (25–50 mm in length, 10–15 mm in diameter) that can be assembled in a suitable preform for optical fibre production. In particular, optical fibres with core diameter of approximately 200 μm , without any cladding (as those used in the studies here described) were obtained by drawing the cylinders welded on a couple of silica handles (Fig. 21.1b).

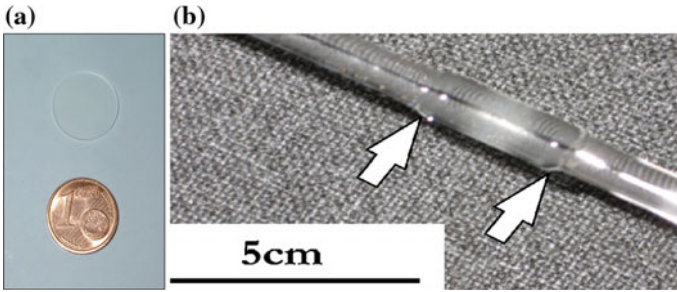
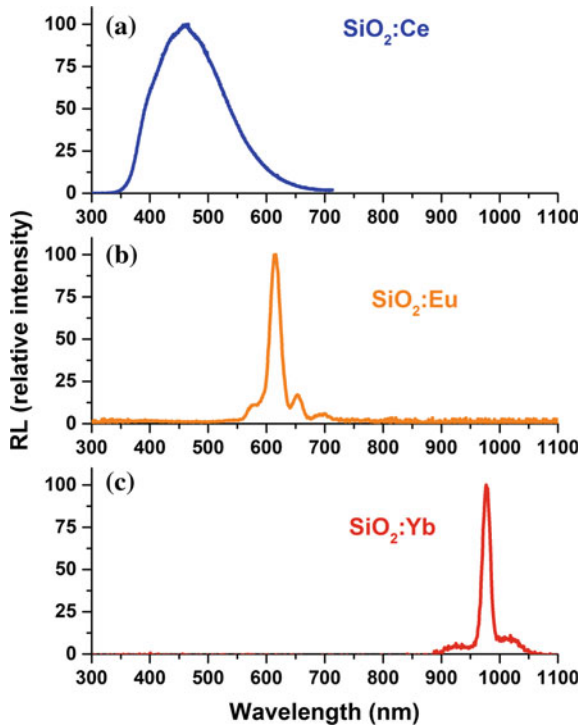


Fig. 21.1 Pictures of scintillating doped silica glass samples in form of bulk (a) and as preform before the fibre drawing (b)

Portions of doped fibres of few mm in length are then cut and connected by fusion splicing to commercial silica optical fibres obtaining the final dosimeters. The RE dopants mainly investigated were cerium, europium and ytterbium. According to the particular RE, the scintillator changed its properties in terms of scintillating time as well as in the RL emission spectrum, as shown in Fig. 21.2.

Fig. 21.2 Radioluminescence spectra of various RE-doped silica optical fibre dosimeters irradiated with 30 kV X-rays



It must be noted that the RL spectra of Fig. 21.2 were obtained by irradiating only the doped fibre with soft X-rays, i.e. below the energy threshold for Cerenkov emission that, for electrons in silica, is approximately equal to 200 keV. Therefore, no contribution of the stem effect occurred.

A brief description of the main features of Ce, Eu and Yb, is reported in the following sections.

21.2.1 *Cerium Doped Silica Optical Fibres*

Traditionally, cerium was used as a dopant because of its very high RL efficiency. Ce-doped silica glasses and fibres have a broad RL spectrum centered at about 450 nm but extending in the whole visible region (Fig. 21.2a). Therefore, the RL signal can be easily detected by means of standard photomultiplier tubes or photodiodes [14, 15]. From the dosimetry point of view the high RL efficiency enables to achieve minimum detectable dose rates much lower than 10^{-2} mGy/s. These systems are also quite fast, with scintillating times of approx. 55 ns due to the radiative transition $5D-4F$ of Ce^{3+} [16].

On the basis of these properties we have seen that Ce-doped fibre are suitable for applications where the stem effect is negligible. In particular, in dosimetry and QA of diagnostic X-rays [17] (taking into account that silica is not a tissue equivalent materials), in brachytherapy [18, 19], as well as in preclinical RT studies.

Moreover, Ce-doped fibres can be used for monitoring ions and protons beams. Indeed particle beams monitoring systems based on these fibres were developed and tested [20]. Furthermore, studies for the use of Ce-doped silica fibres in proton therapy were carried out at the proton therapy centre of the Paul Scherrer Institute with a 138 MeV scanned proton pencil beam. The feature of the system to achieve a real-time measurement of the dose rate enabled to time-resolve different spot scanning sequences of the dose delivering system [21].

The measurement of the depth dose curve of this proton beam in water revealed the presence of a quenching effect of the order of 12% in the region of the Bragg peak, as shown in Fig. 21.3. In fact, the progressive decrease of the detector efficiency with increasing the Linear Energy Transfer (LET), observed in most of the solid state detectors, was expected even in optical fibre dosimeters [5].

21.2.2 *Europium Doped Silica Optical Fibres*

In order to try to face the drawback of the stem effect we investigated a dopant with an emission at higher wavelength than Ce, that is europium. Indeed the RL spectrum of Eu-doped fibre was characterized by a narrow main peak around 620 nm and by weaker peaks below and above this wavelength (see Fig. 21.2b). All these emissions are related to ${}^5D_0 - {}^7F_J$ ($J = 1, 2, 3, 4$) transitions of Eu^{3+} .

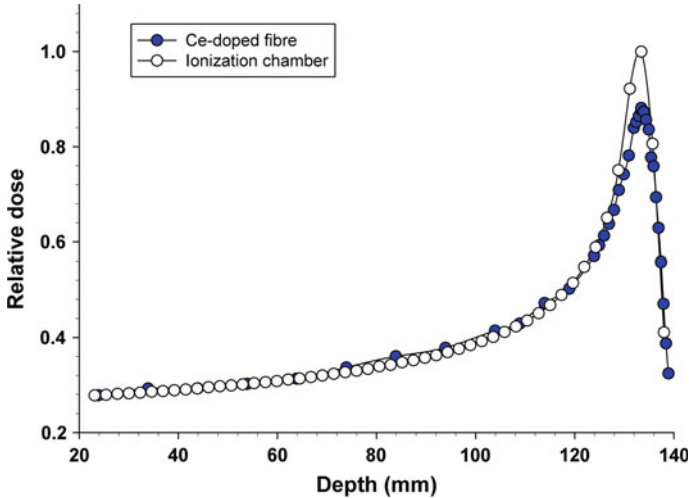


Fig. 21.3 Depth dose curve in water for protons of 138 MeV measured with the Ce-doped fibre and compared with the reference curve obtained with an ionization chamber

However, a background due to stem effect can be observed in the red region as well, in case of unlucky irradiation conditions, like use of large radiation fields or particular orientations of the beam. An example is shown in Fig. 21.4 where the RL emission spectrum of a Eu-doped fibre was measured by irradiating the doped portion with a 6 MeV electron beam at different angles between the beam direction and the fibre axis. A compact UV-VIS spectrometer consisting in a thermoelectric cooled back-thinned CCD array (PrimeTM X, B&WTec Inc., USA) was used for these in situ spectral measurements.

When the angle between the fibre axis and the electron beam axis was equal to 45° , the stem effect was significantly higher than what was observed when the beam interacted orthogonally to the fibre. Indeed, for silica at 6 MeV, the angle at which the Cerenkov light is emitted relative to the direction of the particle is close to 45° . Therefore, at this angle the cone of the Cerenkov light was most efficiently projected into the optical fibre producing a significant increase of the stem effect contribution.

Actually, in this case an effective procedure for removing the stem effect was developed and validated. Indeed, the net area of the Eu^{3+} main peak, i.e., the RL signal over the spurious signal, proved to act as a Cerenkov-free dosimetric signal [7, 8]. Unfortunately, the RL spectra must be collected in order to perform this correction, that is quite complex and time-consuming in the clinical practice.

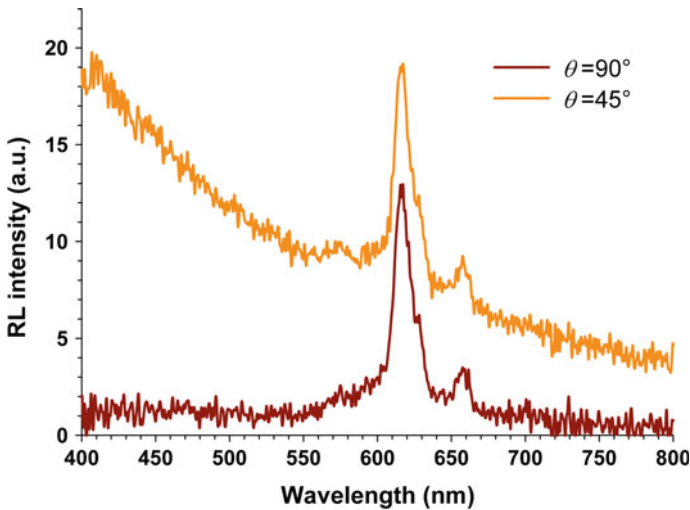


Fig. 21.4 Examples of RL spectra of a Eu-doped fibre irradiated with a 6 MeV electron beam at different angles between the beam direction and the fibre axis

21.2.3 Ytterbium Doped Silica Optical Fibres

The final solution to the problem of the stem effect was achieved by developing Yb-doped silica optical fibres, i.e. exploiting the near infrared emission at 975 nm of Yb^{3+} (see Fig. 21.2c). Indeed, in situ spectral measurements attested that the RL signal of Yb^{3+} remained constant independently of the portion of passive fibre exposed to the radiation beam, and independently of the orientation of beam [22].

Yb-doped silica fibres, like many other scintillators, are characterized by a hysteresis effect, i.e. an increase of the scintillation efficiency with increasing the cumulated dose. This effect is due to the presence of competitive traps that are progressively filled during the irradiation [23, 24]. The study of the nature of these traps by means of thermoluminescence measurements [25] demonstrated that they are deep enough to remain filled at room temperature enabling a good stability and reproducibility of the RL signal after a suitable pre-irradiation [22].

On the basis of these evidences, a prototype of optical detector for real-time measurements of the RL of Yb was implemented. Key element of the system is a thermally cooled silicon avalanche photodiode operating in Geiger mode and characterized by a quantum efficiency of approximately 13% in the wavelength region of interest. Furthermore, a filtering system consisting of two long-pass filters with cut-on wavelength at 950 nm proved to enable the efficient transmission of the Yb-emission and to strongly limit the detection of the stem effect contribution occurring in the lower wavelength region.

Characterization measurements performed with clinical radiotherapy beams attested a satisfactory reproducibility of the system, in addition to the high sensi-

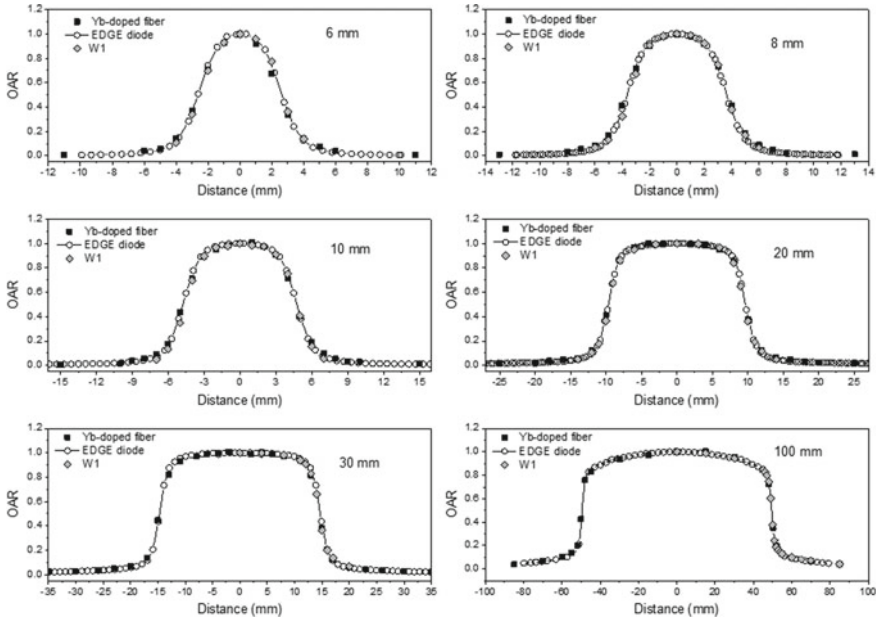


Fig. 21.5 OARs in water of 6 MV FFF X-ray beams of different sizes, measured with the Yb-doped fibre, the commercial scintillator Exradin W1 and the stereotactic diode EDGE detector

tivity, linear dose-rate response, independence of the signal (total counts) of dose rate and impinging beam orientation [26].

Furthermore, the results obtained with Yb-doped silica optical fibre dosimeters proved to be in agreement with those of reference dosimeters in terms of relative dose profiles and output factors [26]. Examples of off-axis ratio (OAR) measurements in water of 6 MV X-rays beam in flattering filter free (FFF) modality, with field size ranging from $6 \times 6 \text{ mm}^2$ to $30 \times 30 \text{ mm}^2$ are shown in Fig. 21.5. In Fig. 21.5 the results of the fibre dosimeters are compared with those obtained by the commercial scintillator system and by a reference diode detector.

21.3 Conclusions

Rare earth doped silica optical fibres showed radioluminescence properties that can be exploited in various dosimetry applications and more in general for ionizing radiation detection and monitoring. For some applications connections with industrial partners have been already established, for other applications, including medical dosimetry, contacts are currently in progress. In particular, the accuracy and precision achieved by Yb-doped fibres in relative dose assessments makes the device promising for QA and in vivo dosimetry studies of modern radiotherapy treatments.

References

1. T. Kron, J. Lehmann, P.B. Greer, Dosimetry of ionising radiation in modern radiation oncology. *Phys. Med. Biol.* **61**, R167–R205 (2016)
2. I. Veronese, E. De Martin, A.S. Martinotti et al., Multi-institutional application of failure mode and effects analysis (FMEA) to CyberKnife stereotactic body radiation therapy (SBRT). *Radiat. Oncol.* **10**, 132 (2015)
3. S. Beddar, L. Beaulieu, *Scintillation Dosimetry* (CRC Press Taylor & Francis Group, Boca Raton, 2016)
4. I.J. Das, J. Morales, P. Francescon, Small field dosimetry: what have we learn? *AIP Conf. Proc.* **1747**, 060001 (2016)
5. A.S. Beddar, T.R. Mackie, F.H. Attix, Cerenkov light generated in optical fibres and other light pipes irradiated by electron beams. *Phys. Med. Biol.* **37**, 925–935 (1992)
6. P. Carrasco, N. Jomet, O. Jordi et al., Characterization of the Exradin W1 scintillator for use in radiotherapy. *Med. Phys.* **42**, 297–304 (2015)
7. I. Veronese, M.C. Cantone, M. Catalano et al., Study of the radioluminescence spectra of doped silica optical fibre dosimeters for stem effect removal. *J. Phys. D Appl. Phys.* **46**, 015101 (2013)
8. I. Veronese, M.C. Cantone, N. Chiodini et al., The influence of the stem effect in Eu-doped silica optical fibres. *Radiat. Meas.* **56**, 316–319 (2013)
9. P.Z.Y. Liu, N. Suchowerska, J. Lambert et al., Plastic scintillation dosimetry: comparison of three solutions for the Cerenkov challenge. *Phys. Med. Biol.* **56**, 5805–5821 (2011)
10. M.A. Clift, P.N. Johnston, D.V. Webb, A temporal method of avoiding the Cerenkov radiation generated in organic scintillator dosimeters by pulsed mega-voltage electron and photon beams. *Phys. Med. Biol.* **47**, 1421–1433 (2002)
11. A.M. Frelin, J.M. Fontbonne, G. Ban et al., Spectral discrimination of Cerenkov radiation in scintillating dosimeters. *Med. Phys.* **32**, 3000–3006 (2005)
12. P. Papaconstadopoulos, L. Archambault, J. Seuntjens, Experimental investigation on the accuracy of plastic scintillators and of the spectrum discrimination method in small photon fields. *Med. Phys.* **44**, 654–664 (2017)
13. N. Chiodini, A. Vedda, I. Veronese, Rare earth doped silica optical fibre sensors for dosimetry in medical and technical applications. *Adv. Opt.* **2014**, 974584 (2014)
14. E. Mones, I. Veronese, F. Moretti et al., Feasibility study for the use of Ce³⁺-doped optical fibres in radiotherapy. *Nucl. Instrum. Methods Phys. Res. A* **562**, 449–455 (2006)
15. E. Mones, I. Veronese, A. Vedda et al., Ce-doped optical fibre as radioluminescent detector in radiotherapy. *Radiat. Meas.* **43**, 888–892 (2008)
16. A. Vedda, N. Chiodini, D. Di Martino et al., Ce³⁺-doped optical fibres for remote radiation dosimetry. *Appl. Phys. Lett.* **85**, 6536 (2004)
17. N. Caretto, N. Chiodini, F. Moretti et al., Feasibility of dose assessment in radiological diagnostic equipments using Ce-doped radio-luminescent optical fibres. *Nucl. Instrum. Methods Phys. Res. A* **612**, 407–411 (2010)
18. M. Carrara, C. Cavatorta, M. Borroni et al., Characterization of a Ce³⁺ doped SiO₂ optical dosimeter for dose measurements in HDR brachytherapy. *Radiat. Meas.* **56**, 312–315 (2013)
19. M. Carrara, C. Tenconi, G. Rossi et al., Temperature dependence of a Ce³⁺ doped SiO₂ radioluminescent dosimeter for in vivo dose measurements in HDR brachytherapy. *Radiat. Meas.* **71**, 324–328 (2014)
20. M. Auger, S. Braccini, T.S. Carzaniga et al., UniBEaM: a silica fibre monitor for charged particle beams. *AIP Conf. Proc.* **1845**, 020015 (2017)
21. I. Veronese, M.C. Cantone, N. Chiodini et al., Feasibility study for the use of cerium-doped silica fibres in proton therapy. *Radiat. Meas.* **45**, 635–639 (2010)
22. I. Veronese, C. De Mattina, M. Fasoli et al., Infrared luminescence for real time ionizing radiation detection. *Appl. Phys. Lett.* **105**, 061103 (2014)
23. I. Veronese, M. Fasoli, M. Martini et al., Phosphorescence of SiO₂ optical fibres doped with Ce³⁺ ions. *Phys. Status Solidi C* **4**, 1024–1027 (2007)

24. F. Moretti, G. Patton, A. Belsky et al., Radioluminescence sensitization in scintillators and phosphors: trap engineering and modeling. *J. Phys. Chem. C* **118**, 9670–9676 (2014)
25. I. Veronese, C. De Mattia, M. Fasoli et al., Role of optical fibre drawing in radioluminescence hysteresis of Yb-doped silica. *J. Phys. Chem. C* **119**, 15572–15578 (2015)
26. I. Veronese, N. Chiodini, S. Cialdi et al., Real-time dosimetry with Yb-doped silica optical fibres. *Phys. Med. Biol.* **62**, 4218–4236 (2017)

Chapter 22

R&D on Electronic Devices and Circuits for the HL-LHC



Alberto Stabile, Attilio Andreazza, Mauro Citterio, Luca Frontini, Valentino Liberali, Chiara Meroni and Jafar Shojaii

Abstract The paper presents the research activities in microelectronics, aiming at improving detection capabilities of future High Energy Physics (HEP) experiments. The output of this research will be the development of novel integrated circuits, to enhance the performance of electronic systems for the High Luminosity Large Hadron Collider (HL-LHC). In particular, the main research activities are focused on monolithic pixel arrays, on new digital architectures for pixel readout in 65 nm CMOS, and on associative memories for several interdisciplinary applications, such as fast tracking for trigger, DNA sequencing, magnetic resonance and image analysis.

22.1 Introduction

The High Luminosity upgrade of the Large Hadron Collider (HL-LHC) is currently expected to begin operations in the second half of 2026, with a nominal instantaneous luminosity of $5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, which corresponds roughly to an average number of 140 inelastic p-p collisions for each beam-crossing, and delivering an integrated luminosity of around 250 fb^{-1} per year of operation. For this reason, electronic systems at LHC must be upgraded to meet these challenging requirements.

Our research activity consists in developing novel integrated circuits, to enhance the performance of electronic systems employed in future detectors at HL-LHC. In particular, we are working on three main research areas.

1. Monolithic pixel arrays in high voltage and high resistivity silicon technology are being designed, and will be characterized under radiation. Currently,

A. Stabile (✉) · A. Andreazza · L. Frontini · V. Liberali
Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano,
Via Celoria 16, 20133 Milan, Italy
e-mail: alberto.stabile@unimi.it

A. Stabile · A. Andreazza · M. Citterio · L. Frontini · V. Liberali · C. Meroni · J. Shojaii
Istituto Nazionale di Fisica Nucleare, Via Celoria 16, 20133 Milan, Italy

J. Shojaii
The University of Melbourne, Swanston St & Tin Alley, Parkville, VIC 3010, Australia

the collaboration is working to find implementations meeting the requirements for deployment at the HL-LHC. Technologies under studies include: AMS 180 nm, LFoundry 150 nm, STMicroelectronic 160 nm, and TowerJazz 180 nm. The microelectronics group in Milano contributes to the prototypes with design of analog and digital blocks, with simulations in different technologies, with irradiations and laboratory characterization.

2. Since 2014, new digital architectures are being designed and implemented in a 65 nm CMOS pixel readout ASIC, aiming at improving radiation hardness through suitable design techniques. This work had lead to a demonstrator within the RD53 collaboration. Milano group will continue to design and characterization of digital blocks (i.e., SRAM, rad-hard logic cells) and will also contribute to the characterization of the RD53 demonstrator.
3. Since 2010, the Milano group is designing and characterizing integrated circuits able to increase the efficiency of LHC trigger systems. The group has designed and characterized four successful Associative Memory ASICs (one of that is now in production for the definitive installation in ATLAS for the 2018). Moreover, since 2015, R&D projects are ongoing, to port the chip design to a 28 nm CMOS technology for the HL-LHC. The new chip will include innovative features to reduce power consumption and silicon area. This chip will be the core of future trigger system for tracking recognition (both at Level-1 and at Level-2). Moreover, it could be useful also to improve pattern-recognition capability in several interdisciplinary applications, such as the DNA sequencing, magnetic resonance image analysis, etc.

All the three areas require a strong microelectronics group, capable to explore new solutions suitable for modern fabrication technologies, to improve the performance of future detectors at HL-LHC. To optimize human, CAD and instrumentation resources, techniques and ideas will be shared.

22.2 Monolithic Pixel Arrays

In high-rate particle-physics applications, a conventional approach is to use *hybrid detectors* integrating two silicon dies: (1) the detector and (2) front-end electronics [1]. An alternative approach consists in using a sole die including all the needed devices [2]. To achieve this task commercial technologies are used. Advantages are a large volume, low cost productions. However, commercial technologies usually have a low resistivity substrate ($1\text{--}10\ \Omega\ \text{cm}$) not enough to generate a large depletion region. Charge collection is obtained by diffusion and charge carriers diffuse in the substrate until either recombine or are collected by an electrode. In addition a long collection time is needed and not all the charges have been collected. For this reason, this devices are not suitable for high rate application and are not still enough radiation hard.

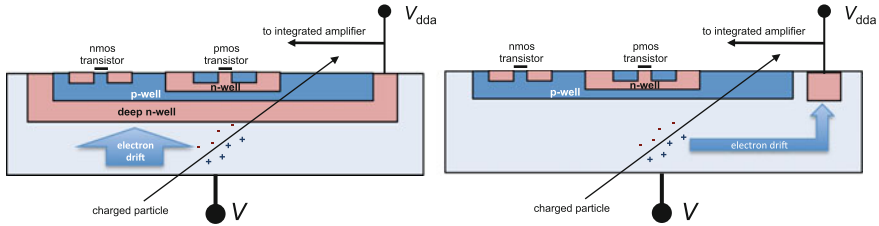


Fig. 22.1 Sketch of monolithic depleted CMOS sensors: large fill-factor design (left) and small fill-factor design (right)

The goal of this research is to design and characterize chip fabricated in technologies able to improve the radiation resistance and timing performance by adding a depleted region below the collection electrode [3]. The depth of the depleted region d is provided by the equation $d = \sqrt{\epsilon\mu\rho(V - V_{bi})}$, where ϵ , μ and ρ are, respectively, the silicon dielectric constant (≈ 1 pF/cm), carrier mobility and resistivity, V_{bi} the junction built-in potential, and V the bias voltage. For this reason, the feasibility of this approach has become possible thanks to: the availability of processes with high-voltage capability, driven by automotive and power management CMOS technologies; foundries accepting wafers or epitaxial substrates with mid-high resistivity ($\rho = 10 - 10000 \Omega \text{ cm}$); the technology nodes in the range from 130 to 180 nm.

A multiple-well process is needed to decouple front-end electronics from the sensitive region. Two different kind of designs are displayed in Fig. 22.1. The left panel of Fig. 22.1 illustrated the “large electrode” design: it uses a deep n-well or a buried layer as a collecting electrode and integrates the front-end electronics inside this well. It provides a short drift path and uniform charge collection efficiency, but it shows an input capacitance of the order of 100 fF. The “small electrode” design is shown in the right panel of Fig. 22.1. It is characterized by a small size collecting electrode, with input capacitance of few fF, and front-end electronics built outside the electrode. The small input capacitance is beneficial to the noise and to the risetime. The drawback is the non-uniformity of the drift field, with longer drift paths, depending on the particle crossing point, and therefore position dependent trapping probability after irradiation.

Depleted-CMOS sensors are undergoing an active research and development phase. Many progresses have been achieved on the monolithic sensors fabricated in CMOS technologies. A good understanding of the depletion zone and charge collection has been achieved and the efficiency is greater than 99% even after irradiation at fluences greater than $10^{15} \text{ n}_{eq}/\text{cm}^2$.

The Milano group has produced “large electrode” sensors in the BCD8 process by STMicroelectronics [4]. Observed performance [5] are in line with detectors in other technologies that achieved and the efficiency is greater than 99% even after irradiation at fluences greater than $10^{15} \text{ n}_{eq}/\text{cm}^2$ [6].

The major missing next step is the integration of the sensors with a high rate readout architecture. Within the ATLAS ITk Collaboration a number of prototypes

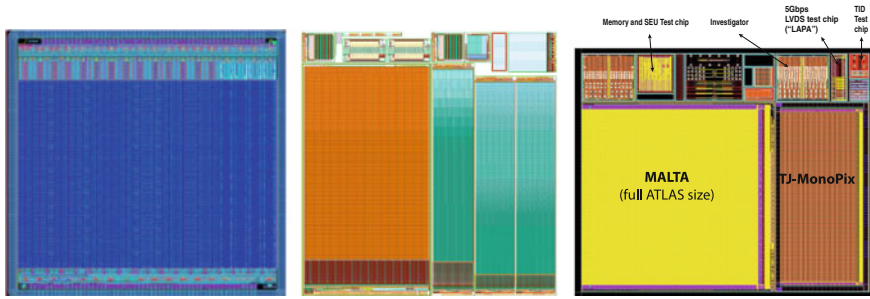


Fig. 22.2 Sketches of some prototype radiation hard monolithic CMOS detectors: LFoundry Monopix (left), AMS MUPIX8 and ATLASPix (middle), and TowerJazz MALTA and Monopix (right)

have been submitted for fabrication. They are shown in Fig. 22.2: the LFoundry 150 nm and AMS aH18 processes are used to implement large electrode designs, while the devices in the TowerJazz 180 nm process use small electrodes. They will allow to evaluate the radiation hardness of the small electrode design and the tolerance to cross-coupling between the readout electronics and the sensor.

22.3 IP Blocks for RD53

Within the RD53 project, we designed a static-RAM (SRAM) cell in 65 nm CMOS technology for high energy physics experiments [7]. A good level of radiation hardness against cumulative dose effects can easily be achieved by using a 65 nm technology, as the oxide thickness is thin enough to provide intrinsic radiation hardness against total ionizing dose.

Single Event Effects (SEEs) can be mitigated by using ad-hoc techniques. The SRAM is based on Dual Interlocked Cells (DICE) to reduce the upset occurrence in the internal nodes of memory latches. Duplicated transistors of the cell are placed at a distance which prevents upsets of multiple nodes within the same cell. Electrical simulations demonstrate a very good tolerance to SEE. The SEEs have been simulated by using fault injection techniques and simulation results demonstrate an increase of tolerance to SEE.

The robustness against Single Event Latch-up (SEL) is achieved by using guard rings that completely surround NMOS transistors in the substrate and PMOS transistors inside n-wells [8].

22.4 Associative Memories

A common problem in the HEP experiments is the particle tracking. Theoretically, this pattern recognition problem could be solved with the help of CPU or GPU. However, for some experiments (such as ATLAS), CPU or GPU are not suitable due the huge amount of trajectories generated at high luminosity. Hence, costs and power consumption exceed a reasonable budget and a dedicated solution has been proposed with the FastTracKer (FTK) project.

Figure 22.3 shows a block diagram of the whole FastTracKer (FTK) system. It stores one billion (10^9) patterns subdivided in 8 Mpatterns per board (128 boards), 128 kpatterns per chip (64 chips per board), and a pattern is composed by 18 bits \times 8 words. The goal is to optimize the pattern density with the design, test and fabrication of large silicon ASICs. Since a large number of input data have to be spread in all chips a I/O signal congestion exists at board level. The solution consists in using LVDS 2 Gbit/s serial links. The entire system have a maximum power of 250 W per AM board. This power is limited by cooling because we are fitting 8192 AMchips in 8 VME crates [9].

The core of FTK is the Associative Memory (AM) chip, a dedicated ASIC able to perform massively parallel pattern-recognition tasks. Especially it performs about one petacomparisons per second. The comparison is bit-wise. Figure 22.4 shows the architecture of a memory block. For each bus (column) and for each pattern (row)

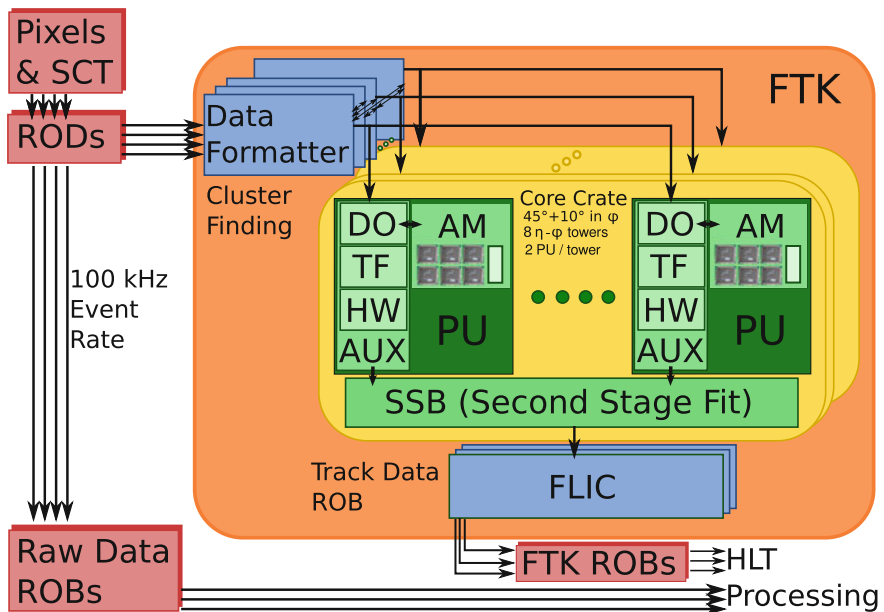


Fig. 22.3 Block diagram of the Fast TracKer

there is a small Content Addressable Memory (CAM) cell array (Fig. 22.4—the little yellow boxes). It compares its own content with all input data received. If it matches a memory latch (Fig. 22.4—the little red boxes) is set to an high logic value. The partial matches are analyzed by “Quorum” logic (Fig. 22.4—the green boxes). If the count is greater than a programmed threshold, the address is put on the output bus by means of a priority readout tree (Fig. 22.4—the yellow boxes).

The AM and commercial CAMs differs substantially because the AM provides a unique capability: the possibility to store partial matches and to find correlations at different times.

In addition, the AM implement an elegant solution to optimize the number of stored patter inside the bank, and the geometrical efficiency. This solution has been called “variable resolution”. It consists in performing pattern matching at reduced and variable resolution first, and then to refine matching resolution using a FPGA. A “don’t care” bit is used to increase the pattern recognition efficiency at different resolutions. Figure 22.5 shows an example of “variable resolution” implementation. With this approach the number of fakes also decrease lowering the whole system power consumption.

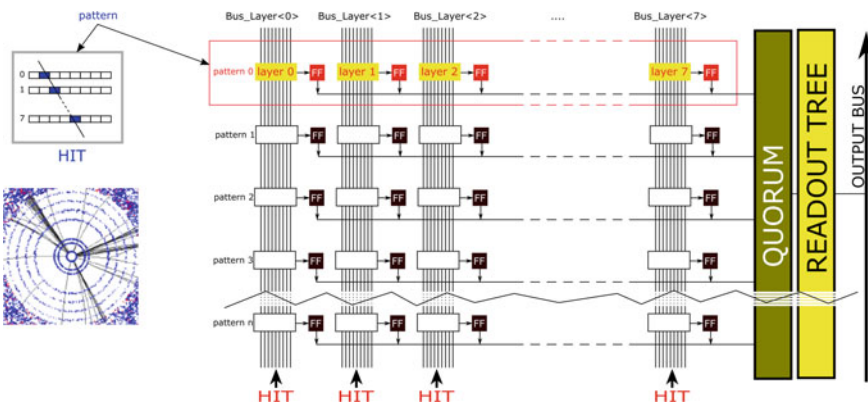


Fig. 22.4 CAM array architecture

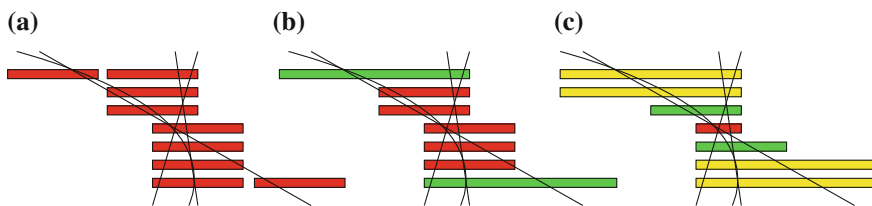


Fig. 22.5 Qualitative example of variable resolution technique: **a** fixed resolution: four patterns to be stored; **b** variable resolution with one don't care bit: one pattern to be stored; **c** variable resolution with improved efficiency and three don't care bits: one pattern to be stored

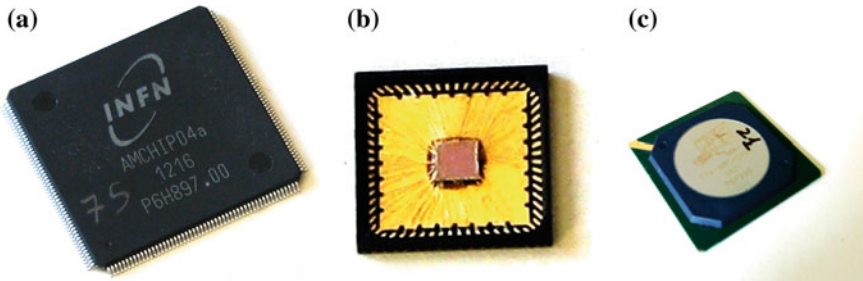
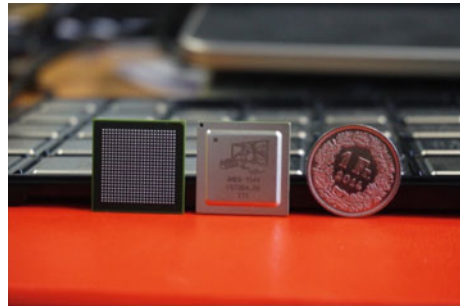


Fig. 22.6 AM chips: **a** version 4; **b** version 5—mini@sic; **c** version 5—MPW

Fig. 22.7 AM chip
(version 6)



In the next future—at the so called phase-II of LHC—the luminosity of LHC will strongly increase. For this reason a more efficient pattern recognition system has to be designed and commissioned. This big challenges will be bring the pileup of the system to an average value of 140 with a peak of 200. The Track Triggers (TTs) are a crucial piece of the phase-II upgrade plan. The AMchips will be used for the level-1 and -2 ATLAS trigger system, and perhaps also the level-1 of CMS experiment will mount AMchips.

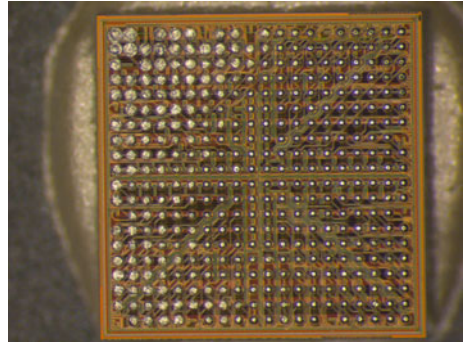
The goal is the evolution of the system design toward the phase-II environments that will be composed by a novel version of AM chips and ATCA racks.

Since 2010 the Milano group design AM chips: the first prototype was the AM04 designed in 65 nm with an area of 14 mm^2 , a cost of 50 k€, and a depth of 8 kpatterns, this chip was fully functional and characterized in 2012 (Fig. 22.6a).

After that, two prototypes was design and characterized in 2012: (1) **AM05-mini@sic** with an area of 4 mm^2 , a cost of 20 k€, and a depth of 256 pattern (Fig. 22.6b); (2) **AM05-MPW** with an area of 12 mm^2 , a cost of 75 k€, and a depth of 5 kpatterns. These prototype had the goal of characterize the LVDS serializer/deserializer and the novel XORAM cells [10] (Fig. 22.6c).

The definite chip for the FTK project was developed in 2014 with and area of 160 mm^2 , a cost of 550 k€, and a depth of 128 kpatterns. It contained about 421 millions of transistors (Fig. 22.7). This chip is comparable in terms of number of transistors to an Intel core Duo.

Fig. 22.8 AM chip
(version 7)



In 2015, we started designing ASICs in a the 28 nm technology. The AM07 has been design for different applications, such as, image analysis, DNA sequencing and Trigger DAQ. The chip area is 10 mm^2 , with a cost of 200 k€ and a pattern density of 16 kpatterns (Fig. 22.8).

In the next future we will design AM08 and AM09. The workflow toward the final ASIC is organized in the following three steps:

- **AM08 prototype:** small area MPW prototype to test all the full custom features, the VHDL logic and the I/O. This chip must be fully functional with smaller memory area than the final ASIC;
- **AM09pre pre-production:** full area ASIC to be fabricated with a full-mask set pilot run. Production corner wafers will be created;
- **AM09 production:** full area ASIC with refinements for the mass production. This chip will comparable in terms of number of transistors to an Quad-core + GPUCore i7 Haswell and about 30 peta comparisons per second per chip.

The AM09 will be developed built on the AM08 extending the memory area, therefore the specification of both versions must be compatible.

AM08 and AM09 will contain 16 kpatterns and 3×128 kpattern low-power CMOS associative memories, organized into 4 (AM08) and 96 (AM09) cores respectively. It is fabricated using very high performance, high reliability CMOS technology at 28 nm). Its standby current is stable within the range of operating temperature. The AM08 and AM09 devices are well designed for high energy applications, and particularly well suited for ATLAS and CMS trigger applications. The AM08 and AM09 operate from a single core power supply of 0.9–1.1 V and all data inputs and outputs are fully LVDS18 compatible. The LVDS18 I/O run at up to 2 Gbit/s.

22.5 Conclusion

The construction and operation of tracking detector at the High Luminosity LHC will require new technological solutions for sensors, front-end electronics and trigger systems. The Milano group active on many of these developments.

Radiation hard CMOS sensors have been realized using technologies with high-voltage capabilities on high resistivity substrate and feature sizes in the range 150–180 μm . They provide an interesting opportunity for future tracking detectors because they allow the integration of the sensitive volume with the front-end electronics. The construction of monolithic devices is now being explored with several foundries.

References

1. G. Aad et al., JINST **3**, P07007 (2008). <https://doi.org/10.1088/1748-0221/3/07/P07007>; CMS Collaboration, JINST **3**, S08004 (2008). <https://doi.org/10.1088/1748-0221/3/08/S08004>; LHCb collaboration, LHCb VELO upgrade technical design report, CERN-LHCC-2013-021, <https://cds.cern.ch/record/1624070>
2. G. Contin et al., JINST **10**, C03026 (2015). <https://doi.org/10.1088/1748-0221/10/03/C03026>; ALICE Collaboration, J. Phys. G **41**, 087002 (2014). <https://doi.org/10.1088/0954-3899/41/8/087002>; T. Behnke et al., *International Linear Collider Technical Design Report—Volume 4: Detectors*, [arXiv:1306.6329](https://arxiv.org/abs/1306.6329)
3. I. Peric, Nucl. Instr. Method A **582**, 876 (2007). <https://doi.org/10.1016/j.nima.2007.07.108>
4. R. Roggero et al., *Proceedings of the 25th International Symposium on Power Semiconductor Devices and ICs (ISPSD2013)*, May 2013, pp. 361–364
5. A. Andreazza et al., JINST **11**, C11038 (2016). <https://doi.org/10.1088/1748-0221/11/11/C11038>
6. I. Peric et al., JINST **12**, C02030 (2017). <https://doi.org/10.1088/1748-0221/12/02/C02030>
7. N. Demaria et al., CHIPIX65: Developments on a new generation pixel readout ASIC in CMOS 65 nm for HEP experiments, in *Proceedings of the IEEE International Workshop on Advances in Sensors and Interfaces (IWASI)*, June 2015. <https://doi.org/10.1109/IWASI.2015.7184947>
8. S. Shojaii et al., A radiation hardened static RAM for high-energy physics experiments, in *Proceedings of the International Conference on Microelectronics (MIEL)*, May 2014. <https://doi.org/10.1109/MIEL.2014.6842164>
9. A. Andreani et al., The AMchip04 and the processing unit prototype for the FastTracker. IOP J. Instr. **7**, C08007 (2012). <https://doi.org/10.1088/1748-0221/7/08/C08007>
10. L. Frontini et al., A new XOR-based content addressable memory architecture, in *Proceedings of the ICECS, Seville* (2012), pp. 701–704. <https://doi.org/10.1109/ICECS.2012.6463629>

Chapter 23

Precise Perturbative QCD Predictions for Large Hadron Collider Physics



Giancarlo Ferrera

Abstract We briefly summarize some theoretical advances on perturbative QCD calculations relevant for the physics of the Large Hadron Collider (LHC). As an explicit example, we discuss the next-to-next-to-leading order (NNLO) QCD calculation for the production of a Higgs boson decaying to bottom quarks in association with a vector boson. We show illustrative numerical results for cross sections and associated distributions with typical kinematical cuts applied in the LHC experimental analysis.

23.1 Introduction

The successful operation of the Large Hadron Collider (LHC) at CERN during its first run (2010–2013) at a centre-of-mass energy (\sqrt{s}) of 7/8 TeV led to several outstanding physics results and in particular to the discovery of the Higgs boson particle [1, 2]. At the current level of experimental precision all the measurements concerning this particle are compatible with the Standard Model (SM) predictions but substantial deviations from the SM predictions are still possible. Currently, in the second run (2015–2018), the LHC is exploring the higher energy regime at a center-of-mass energy of 13 TeV. Looking to the future, the LHC will operate at $\sqrt{s} = 14$ TeV in the third run (2020–2022) and, after a significant upgrade of the luminosity performance by a factor 10, the High-Luminosity Large Hadron Collider (HL-LHC) should be operational after 2025.

G. Ferrera (✉)

Dipartimento di Fisica “Aldo Pontremoli”, Università di Milano, Milan, Italy

e-mail: giancarlo.ferrera@mi.infn.it

URL: <http://pcteserver.mi.infn.it/~ferrera/>

G. Ferrera

INFN, Sezione di Milano, 20133 Milan, Italy

© Springer Nature Switzerland AG 2018

P. F. Bortignon et al. (eds.), *Toward a Science Campus in Milan*,

https://doi.org/10.1007/978-3-030-01629-6_23

In the next 20 years the LHC will therefore provide a large amount of high precision data in an unexplored region of very high energies. This will allow us to perform precision tests of the Standard Model with detailed studies of its particles and to observe possible effects of physics beyond the SM.

Among various future measurements, the detailed study of the properties of the recent discovered Higgs particle will be of primary importance [3]. Detailed comparisons of theory and experiment can either confirm the SM predictions or indicate the presence of new physics effects. To this aim it will be essential to measure the processes which give information on Higgs boson couplings to other SM particles and to compare the measurements with the most accurate theoretical predictions.

23.2 Perturbative QCD Calculations

In order to fully exploit the physics information contained in the high precision data and the discovery potential of the LHC, it is important to provide theoretical predictions at a comparable level of accuracy.

At hadronic machines and in particular at the LHC, all the interesting reactions at high transverse momenta involve a hard scattering of partons inside hadrons. Precise theoretical predictions thus require a good control of the corresponding QCD cross sections. The standard framework to perform QCD predictions at hadron colliders is provided by the factorization formula. According to it, the cross section $\sigma(Q^2)$ for a generic scattering process characterized by a hard scale Q can be written in the following form:

$$\sigma(Q^2) = \sum_{a,b} \int_0^1 dx_1 \int_0^1 dx_2 f_{a/h_1}(x_1, \mu_F^2) f_{b/h_2}(x_2, \mu_F^2) \hat{\sigma}_{ab}(\hat{s}, Q^2, \mu_F^2) + \mathcal{O}((\Lambda/Q)^p), \quad (23.1)$$

where $f_{a/h}(x, \mu_F)$ are universal non-perturbative parton distribution functions (PDFs) of the colliding hadrons, which give the distribution of the parton a inside the hadron h at the factorization scale μ_F as a function of the longitudinal momentum fraction x carried by the parton, $\hat{\sigma}_{ab}$ are the process-dependent partonic cross sections and $\hat{s} = x_1 x_2 s$ is the partonic centre-of-mass energy squared. The hard scale of the process Q is typically set by the transverse momentum or the invariant mass of the produced (and triggered) final state system, and it has to be much larger of the QCD scale, $Q \gg \Lambda$. Finally the term $\mathcal{O}((\Lambda/Q)^p)$ (with $p \geq 1$) represents non-perturbative power suppressed contributions to the cross section. A generic infrared and collinear safe partonic cross section $\hat{\sigma}_{ab}$ ¹ can be reliably calculated in perturbation theory through a power expansion in the strong coupling α_s

¹That is an observable independent of the number of soft and collinear final state particles.

$$\hat{\sigma}_{ab}(\hat{s}, Q^2, \mu_F^2) = \sum_{n=0}^{\infty} \left(\frac{\alpha_S(\mu_R^2)}{\pi} \right)^n \hat{\sigma}_{ab}^{(n)}(\hat{s}, Q^2, \mu_F^2, \mu_R^2), \quad (23.2)$$

where μ_R is the renormalization scale.

In order to provide accurate theoretical predictions through the factorization formula in (23.1), it is essential to have a good knowledge of both the PDFs² and the partonic cross sections, together with a reliable estimate of the corresponding uncertainties. The knowledge of partonic cross sections can be systematically improved by performing the calculation of higher-order perturbative QCD corrections in (23.2).³

Typically lowest order or leading order (LO) QCD calculations give only a rough estimate of the corresponding cross sections. The first reliable estimate can be obtained through the calculation of the next-to-leading order (NLO) corrections, while a precise prediction with a robust estimate of the perturbative uncertainty requires the knowledge of next-to-next-to-leading (NNLO) corrections.

However the computation of the NLO and NNLO radiative corrections is particularly involved due to the presence, in the intermediate steps of the calculation, of infrared (IR) divergences which prevent the straightforward use of numerical techniques and require specific analytic treatment. At NLO the structure of IR divergences is well understood and the development of general algorithms made possible to perform NLO calculation for a wide variety of the process [5, 6]. At NNLO a general algorithm to perform fully differential calculations is still lacking. However in the last few years new ideas were developed [7–15] and fully differential NNLO QCD calculation for an increasing number of processes were recently performed [16–28].

23.3 Associated VH Production at the LHC

As an explicit example in this section we consider the inclusive hard scattering reaction

$$h_1 + h_2 \rightarrow VH + X \rightarrow l_1 l_2 b \bar{b} + X \quad (23.3)$$

where the collision of the hadrons (protons in the case of the LHC) h_1 and h_2 produces the system VH ($V = W^\pm, Z$) which decays in turn into a lepton pair $l_1 l_2$ and a bottom quark-antiquark pair $b \bar{b}$, while we denote with X the accompanying final state QCD radiation. The hard-scattering scale of the process is set by the invariant mass M_{VH} .

Within the narrow width approximation for the Higgs boson, ($\Gamma_H \ll m_H$), the perturbative QCD expansion of the cross section up to complete NNLO can be written as

²For a recent review about the state of the art on PDF determination see [4].

³Given the larger value of the QCD coupling compared to the electroweak one, at typical momentum scales of LHC processes, the impact of the electroweak corrections is in general less relevant.

Table 23.1 Cross sections and their numerical uncertainties for $pp \rightarrow W^+H + X \rightarrow l\nu b\bar{b} + X$ at LHC with $\sqrt{s} = 13$ TeV. The applied kinematical cuts are described in the text

σ (fb)	NNLO (prod) + NLO (dec)	Full NNLO
W^+H	3.94 ± 0.02	3.70 ± 0.02

$$\begin{aligned}
d\sigma_{h_1 h_2 \rightarrow VH \rightarrow V b\bar{b}}^{\text{NNLO}} = & \left[d\sigma_{h_1 h_2 \rightarrow VH}^{(0)} \times \frac{d\Gamma_{H \rightarrow b\bar{b}}^{(0)} + d\Gamma_{H \rightarrow b\bar{b}}^{(1)} + d\Gamma_{H \rightarrow b\bar{b}}^{(2)}}{\Gamma_{H \rightarrow b\bar{b}}^{(0)} + \Gamma_{H \rightarrow b\bar{b}}^{(1)} + \Gamma_{H \rightarrow b\bar{b}}^{(2)}} \right. \\
& + d\sigma_{h_1 h_2 \rightarrow VH}^{(1)} \times \frac{d\Gamma_{H \rightarrow b\bar{b}}^{(0)} + d\Gamma_{H \rightarrow b\bar{b}}^{(1)}}{\Gamma_{H \rightarrow b\bar{b}}^{(0)} + \Gamma_{H \rightarrow b\bar{b}}^{(1)}} \\
& \left. + d\sigma_{h_1 h_2 \rightarrow VH}^{(2)} \times \frac{d\Gamma_{H \rightarrow b\bar{b}}^{(0)}}{\Gamma_{H \rightarrow b\bar{b}}^{(0)}} \right] \times \text{Br}(H \rightarrow b\bar{b}), \quad (23.4)
\end{aligned}$$

where $\Gamma_{H \rightarrow b\bar{b}}$ and Γ_H are the Higgs boson partial decay width to bottom quarks and the total decay width respectively and $\text{Br}(H \rightarrow b\bar{b}) = \Gamma_{H \rightarrow b\bar{b}}/\Gamma_H$ is the corresponding branching ratio.

The formula in (23.4) has been implemented in [28] by using the q_T subtraction formalism [10, 29] for the VH production corrections [17, 20] and the *CoLoR-FulNNLO* formalism [8, 9, 30] for the $H \rightarrow b\bar{b}$ decay [31].

We now present illustrative numerical results for the LHC at $\sqrt{s} = 13$ TeV [28]. We use the PDF4LHC PDFs set [32] at NNLO with $\alpha_S(m_Z) = 0.118$ and the G_μ scheme for the electroweak couplings with the following input parameters: $G_F = 1.1663787 \times 10^{-5}$ GeV⁻², $m_Z = 91.1876$ GeV, $m_W = 80.385$ GeV, $\Gamma_Z = 2.4952$ GeV, $\Gamma_W = 2.085$ GeV, $m_t = 172$ GeV, $m_b = 4.18$ GeV, $m_H = 125$ GeV, $\Gamma_H = 4.070$ MeV and $\text{Br}(H \rightarrow b\bar{b}) = 0.578$ [33].

We set the renormalization and factorization scales to $\mu_R = \mu_F = M_{VH}$ and the renormalization scale for the $H \rightarrow b\bar{b}$ coupling to $\mu_r = m_H$.

We consider W^+H production and decay with the following kinematical cuts [34]. The charged lepton has transverse momentum $p_T^l > 15$ GeV and pseudorapidity $|\eta_l| < 2.5$, the missing transverse momentum of the event is required to be $p_T^v > 30$ GeV, the W boson has a transverse momentum $p_T^W > 150$ GeV and we require the presence of at least two b -jets⁴ with $p_T^b > 25$ GeV and $|\eta_b| < 2.5$. In Table 23.1 we report the corresponding cross sections together with an estimate of the numerical uncertainties of our calculation and we compare the *full* NNLO prediction with the *partial* NNLO prediction which contains the NLO corrections for decay. The inclusion of the *full* NNLO corrections reduces the cross section by around 6%.

In Fig. 23.1 (left) we show the transverse-momentum distribution $p_T^{b\bar{b}}$ of the leading b -jet pair. We observe that the *full* NNLO prediction have an substantial effect on the shape of the spectrum. In particular the cross section is increased by around 2–5% for $p_T^{b\bar{b}} \lesssim 140$ GeV and it is decreased by around 6–8% for $p_T^{b\bar{b}} \gtrsim 140$ GeV.

⁴Jets are reconstructed with the flavour- k_T algorithm with $R = 0.5$ [35].

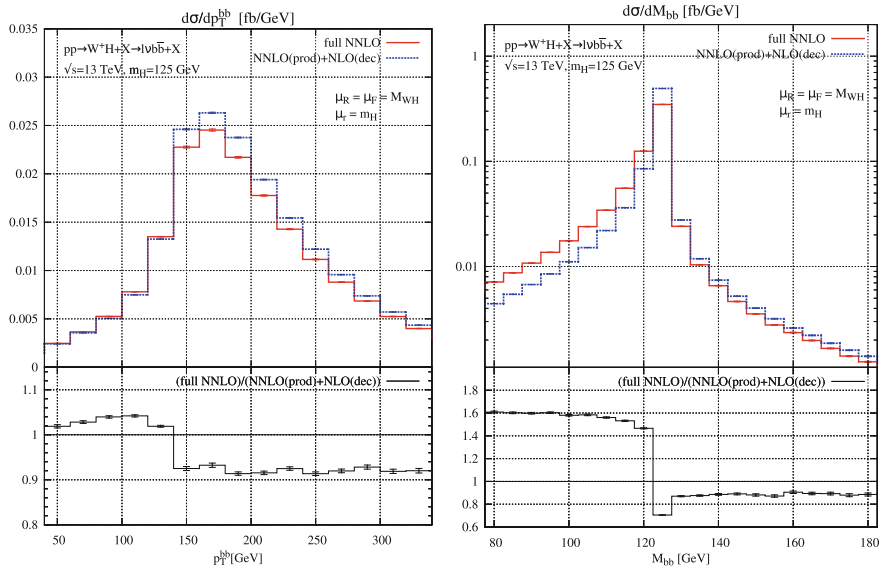


Fig. 23.1 $pp \rightarrow W^+H + X \rightarrow l\nu_l b\bar{b} + X$ at LHC with $\sqrt{s} = 13$ TeV. Transverse-momentum distribution (left panel) and invariant mass distribution (right panel) of the leading b -jet pair computed at full NNLO (red solid) and partial NNLO (blue dashes). The lower panels show the ratios of the results. The applied cuts are described in the text

The invariant mass distribution of the leading b -jet pair, M_{bb} is presented in Fig. 23.1 (right) where we compare the *full* NNLO QCD prediction with the *partial* NNLO one. For this observable the effect of the NNLO corrections to the decay rate are even more important. While the position of the peak is rather stable ($M_{bb} \simeq m_H$), the spectrum receives large (up to +60%) positive corrections for $M_{bb} < m_H$ and substantial (from -30 to -10%) negative corrections for $M_{bb} \gtrsim m_H$.

As it was shown in the case we have explicitly considered, higher-order QCD calculations are essential to obtain precise theoretical predictions for hard-scattering cross sections and associated distributions at the LHC.

23.4 Conclusions

In order to fully exploit the physics information contained in the high precision data of the Large Hadron Collider (LHC), it is essential to provide theoretical predictions at a comparable level of accuracy. We have briefly summarized some theoretical advances on perturbative QCD calculations relevant for the LHC physics. In particular we discussed the As an explicit example, we have discussed the impact of the next-to-next-to-leading order (NNLO) QCD correction for the production and decay of a Higgs boson in association with a vector boson at the LHC.

Acknowledgements We acknowledge the support of Fondazione Cariplo under the grant number 2015-0761.

References

1. G. Aad et al. [ATLAS Collaboration], *Phys. Lett. B* **716**, 1 (2012). <https://doi.org/10.1016/j.physletb.2012.08.020>, [arXiv:1207.7214](https://arxiv.org/abs/1207.7214) [hep-ex]
2. S. Chatrchyan et al. [CMS Collaboration], *Phys. Lett. B* **716**, 30 (2012). <https://doi.org/10.1016/j.physletb.2012.08.021>, [arXiv:1207.7235](https://arxiv.org/abs/1207.7235) [hep-ex]
3. S. Heinemeyer et al. [LHC Higgs Cross Section Working Group]. <https://doi.org/10.5170/CERN-2013-004>, [arXiv:1307.1347](https://arxiv.org/abs/1307.1347) [hep-ph]
4. J. Gao, L. Harland-Lang, J. Rojo, [arXiv:1709.04922](https://arxiv.org/abs/1709.04922) [hep-ph]
5. S. Frixione, Z. Kunszt, A. Signer, *Nucl. Phys. B* **467**, 399 (1996). [https://doi.org/10.1016/0550-3213\(96\)00110-1](https://doi.org/10.1016/0550-3213(96)00110-1) [hep-ph/9512328]
6. S. Catani, M.H. Seymour, *Phys. Lett. B* **378**, 287 (1996); S. Catani, M.H. Seymour, *Nucl. Phys. B* **485**, 291 (1997). [Erratum-ibid. *B* **510**, 503 (1998)]. S. Catani, S. Dittmaier, M.H. Seymour, Z. Trócsányi, *Nucl. Phys. B* **627**, 189 (2002)
7. A. Gehrmann-De Ridder, T. Gehrmann, E.W.N. Glover, *JHEP* **0509**, 056 (2005). <https://doi.org/10.1088/1126-6708/2005/09/056> [hep-ph/0505111]
8. G. Somogyi, Z. Trocsanyi, V. Del Duca, *JHEP* **0701**, 070 (2007). <https://doi.org/10.1088/1126-6708/2007/01/070> [hep-ph/0609042]
9. G. Somogyi, Z. Trocsanyi, *JHEP* **0701**, 052 (2007). <https://doi.org/10.1088/1126-6708/2007/01/052> [hep-ph/0609043]
10. S. Catani, M. Grazzini, *Phys. Rev. Lett.* **98**, 222002 (2007). <https://doi.org/10.1103/PhysRevLett.98.222002> [hep-ph/0703012]
11. M. Czakon, *Phys. Lett. B* **693**, 259 (2010). <https://doi.org/10.1016/j.physletb.2010.08.036>, [arXiv:1005.0274](https://arxiv.org/abs/1005.0274) [hep-ph]
12. R. Boughezal, C. Focke, X. Liu, F. Petriello, *Phys. Rev. Lett.* **115**(6), 062002 (2015). <https://doi.org/10.1103/PhysRevLett.115.062002>, [arXiv:1504.02131](https://arxiv.org/abs/1504.02131) [hep-ph]
13. J. Gaunt, M. Stahlhofen, F.J. Tackmann, J.R. Walsh, *JHEP* **1509**, 058 (2015). [https://doi.org/10.1007/JHEP09\(2015\)058](https://doi.org/10.1007/JHEP09(2015)058), [arXiv:1505.04794](https://arxiv.org/abs/1505.04794) [hep-ph]
14. F. Caola, K. Melnikov, R.Rötsch, *Eur. Phys. J. C* **77**(4), 248 (2017). <https://doi.org/10.1140/epjc/s10052-017-4774-0>, [arXiv:1702.01352](https://arxiv.org/abs/1702.01352) [hep-ph]
15. M. Cacciari, F.A. Dreyer, A. Karlberg, G.P. Salam, G. Zanderighi, *Phys. Rev. Lett.* **115**(8), 082002 (2015). <https://doi.org/10.1103/PhysRevLett.115.082002>, [arXiv:1506.02660](https://arxiv.org/abs/1506.02660) [hep-ph]
16. S. Catani, L. Cieri, D. de Florian, G. Ferrera, M. Grazzini, *Phys. Rev. Lett.* **108**, 072001 (2012). Erratum: [*Phys. Rev. Lett.* **117**(8), 089901 (2016)]. <https://doi.org/10.1103/PhysRevLett.108.072001>. <https://doi.org/10.1103/PhysRevLett.117.089901>, [arXiv:1110.2375](https://arxiv.org/abs/1110.2375) [hep-ph]
17. G. Ferrera, M. Grazzini, F. Tramontano, *Phys. Rev. Lett.* **107**, 152003 (2011). <https://doi.org/10.1103/PhysRevLett.107.152003>, [arXiv:1107.1164](https://arxiv.org/abs/1107.1164) [hep-ph]
18. M. Grazzini, S. Kallweit, D. Rathlev, A. Torre, *Phys. Lett. B* **731**, 204 (2014). <https://doi.org/10.1016/j.physletb.2014.02.037>, [arXiv:1309.7000](https://arxiv.org/abs/1309.7000) [hep-ph]
19. G. Ferrera, M. Grazzini, F. Tramontano, *JHEP* **1404**, 039 (2014). [https://doi.org/10.1007/JHEP04\(2014\)039](https://doi.org/10.1007/JHEP04(2014)039), [arXiv:1312.1669](https://arxiv.org/abs/1312.1669) [hep-ph]
20. G. Ferrera, M. Grazzini, F. Tramontano, *Phys. Lett. B* **740**, 51 (2015). <https://doi.org/10.1016/j.physletb.2014.11.040>, [arXiv:1407.4747](https://arxiv.org/abs/1407.4747) [hep-ph]
21. T. Gehrmann, M. Grazzini, S. Kallweit, P. Maierhöfer, A. von Manteuffel, S. Pozzorini, D. Rathlev, L. Tancredi, *Phys. Rev. Lett.* **113**(21), 212001 (2014). <https://doi.org/10.1103/PhysRevLett.113.212001>, [arXiv:1408.5243](https://arxiv.org/abs/1408.5243) [hep-ph]
22. F. Cascioli et al., *Phys. Lett. B* **735**, 311 (2014). <https://doi.org/10.1016/j.physletb.2014.06.056>, [arXiv:1405.2219](https://arxiv.org/abs/1405.2219) [hep-ph]

23. M. Grazzini, S. Kallweit, D. Rathlev, JHEP **1507**, 085 (2015). [https://doi.org/10.1007/JHEP07\(2015\)085](https://doi.org/10.1007/JHEP07(2015)085), [arXiv:1504.01330](https://arxiv.org/abs/1504.01330) [hep-ph]
24. R. Boughezal, C. Focke, W. Giele, X. Liu, F. Petriello, Phys. Lett. B **748**, 5 (2015). <https://doi.org/10.1016/j.physletb.2015.06.055>, [arXiv:1505.03893](https://arxiv.org/abs/1505.03893) [hep-ph]
25. M. Grazzini, S. Kallweit, D. Rathlev, Phys. Lett. B **750**, 407 (2015). <https://doi.org/10.1016/j.physletb.2015.09.055>, [arXiv:1507.06257](https://arxiv.org/abs/1507.06257) [hep-ph]
26. R. Boughezal, J.M. Campbell, R.K. Ellis, C. Focke, W.T. Giele, X. Liu, F. Petriello, Phys. Rev. Lett. **116**(15), 152001 (2016). <https://doi.org/10.1103/PhysRevLett.116.152001>, [arXiv:1512.01291](https://arxiv.org/abs/1512.01291) [hep-ph]
27. R. Boughezal, J.M. Campbell, R.K. Ellis, C. Focke, W. Giele, X. Liu, F. Petriello, C. Williams, Eur. Phys. J. C **77**(1), 7 (2017). <https://doi.org/10.1140/epjc/s10052-016-4558-y>, [arXiv:1605.08011](https://arxiv.org/abs/1605.08011) [hep-ph]
28. G. Ferrera, G. Somogyi, F. Tramontano, [arXiv:1705.10304](https://arxiv.org/abs/1705.10304) [hep-ph]
29. S. Catani, L. Cieri, D. de Florian, G. Ferrera, M. Grazzini, Nucl. Phys. B **881**, 414 (2014). <https://doi.org/10.1016/j.nuclphysb.2014.02.011>, [arXiv:1311.1654](https://arxiv.org/abs/1311.1654) [hep-ph]
30. V. Del Duca, C. Duhr, A. Kardos, G. Somogyi, Z. Szo, Z. Trocsenyi, Z. Tulipant, Phys. Rev. D **94**(7), 074019 (2016). <https://doi.org/10.1103/PhysRevD.94.074019>, [arXiv:1606.03453](https://arxiv.org/abs/1606.03453) [hep-ph]
31. V. Del Duca, C. Duhr, G. Somogyi, F. Tramontano, Z. Trocsenyi, JHEP **1504**, 036 (2015). [https://doi.org/10.1007/JHEP04\(2015\)036](https://doi.org/10.1007/JHEP04(2015)036), [arXiv:1501.07226](https://arxiv.org/abs/1501.07226) [hep-ph]
32. J. Butterworth et al., J. Phys. G **43**, 023001 (2016); S. Dulat et al., Phys. Rev. D **93**(3), 033006 (2016); L.A. Harland-Lang, A.D. Martin, P. Motylinski, R.S. Thorne, Eur. Phys. J. C **75**(5), 204 (2015); R.D. Ball et al. [NNPDF Collaboration], JHEP **1504**, 040 (2015); S. Carrazza, S. Forte, Z. Kassabov, J.I. Latorre, J. Rojo, Eur. Phys. J. C **75**(8), 369 (2015); J. Gao, P. Nadolsky, JHEP **1407**, 035 (2014)
33. S. Dittmaier et al. [LHC Higgs Cross Section Working Group], [arXiv:1101.0593](https://arxiv.org/abs/1101.0593) [hep-ph]
34. D. de Florian et al. [LHC Higgs Cross Section Working Group], [arXiv:1610.07922](https://arxiv.org/abs/1610.07922) [hep-ph]
35. A. Banfi, G.P. Salam, G. Zanderighi, Eur. Phys. J. C **47**, 113 (2006)

Chapter 24

LASA Cs₂Te Photocathodes: The Electron Source for XFELs



Carlo Pagani, Paolo Michelato, Laura Monaco and Daniele Sertore

Abstract Laser triggered photocathodes are key components of the electron sources for the XFELs. Research on semiconductor photocathodes started in the 90s at INFN LASA Milano by studying multialkali antimonied compounds (Cs₃Sb, K₃Sb, and K₂CsSb), showing that the high QE was coupled with very high sensitivity to vacuum pollution. To avoid these deterioration issues, we focused our activity on Cs₂Te which, while sensitive to UV light, has better resistance to gas contamination. Since then, we have a dedicated R&D program aimed at developing a always more reliable photocathode and its associated transport systems. Today our photocathodes are used at FLASH and at European XFEL at DESY Hamburg, PITZ at DESY Zeuthen, APEX in LBNL, FAST at FNAL and at LCLS-II at SLAC.

24.1 Photocathodes for Free Electron Lasers

Photocathodes are key element in the modern high brightness electron sources, necessary to achieve the beam performance required for the operation of FELs. The final wavelength λ_{FEL} achievable by an FEL is given by the relation [7]

$$\frac{\epsilon}{\gamma} < \frac{\lambda_{\text{FEL}}}{4\pi} \quad (24.1)$$

where ϵ is the beam emittance, i.e. the volume occupied by the beam in the six-dimensional phase space and γ the relativistic factor.

The emittance is generally a sum in quadrature of different contributions:

$$\epsilon = \sqrt{\epsilon_{\text{RF}}^2 + \epsilon_{\text{SC}}^2 + \epsilon_{\text{therm}}^2 + \dots} \quad (24.2)$$

C. Pagani (✉)

Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano,
Via Celoria 16, 20133 Milano, Italy
e-mail: carlo.pagani@mi.infn.it

C. Pagani · P. Michelato · L. Monaco · D. Sertore
INFN - LASA, Via fratelli Cervi 201, 20090 Segrate (Milano), Italy

© Springer Nature Switzerland AG 2018

P. F. Bortignon et al. (eds.), *Toward a Science Campus in Milan*,
https://doi.org/10.1007/978-3-030-01629-6_24

where the photocathode related quantity is the thermal emittance ϵ_{therm} . This term depends on the characteristic of the photocathode material and, in case of semiconductor, can be written, following [5], as

$$\epsilon_{\text{therm}} = \sigma_{\text{laser}} \sqrt{\frac{2 E_{\text{Kin}}}{3 m_0 c^2}} \quad (24.3)$$

being σ_{laser} the laser spot size on the photocathode, $m_0 c^2$ the energy of the electron rest mass and $E_{\text{Kin}} = h\nu - (E_G + E_A)$ the kinetic energy of the emitted electron given by the difference between the photon energy and the sum of energy gap (E_G) and electron affinity (E_A).

A proper choice of the photocathode material is of primary importance for achieving the desired machine parameters. While antimony based alkaly photocathodes (i.e. Cs_3Sb , K_2CsSb , NaKC sSb , etc.) are more promising in term of final thermal emittance and have the advantage of being operable in the visible light, alkaly telluride (i.e. K_2Te , Cs_2Te) are more robust in term of acceptable vacuum condition but require a laser operated in the ultra violet windows being not sensitive to visible light.

Besides the thermal emittance, other parameters are important in the selection of a proper photocathodes for their operation in FELs such as:

- Response time: the generated electron bunch needs to be prompt in order to ‘copy’ the laser profile and hence minimize emittance contribution
- Dark current: it is essential to minimize the emitted dark current to reduce activation of components and thermal loads on downstream components
- Uniformity of emitting area: contribution to final emittance from un-uniformity of the emitting area might be a relevant contribution
- Roughness: film roughness contribution to the final electron beam emittance might be a large fraction of the final emittance.

A final parameter that characterize a photocathode is its Quantum Efficiency (QE) given by

$$\text{QE} = \frac{\# \text{ emitted electrons}}{\# \text{ incident photons}} \quad (24.4)$$

Typical values for metal photocathodes are QEs in the 10^{-5} – 10^{-4} range. These photocathodes have fast response time (in the fs range), they are robust w.r.t. to vacuum condition and usually are sensitive in the UV or DUV range. More performing photocathodes are based on semiconductor materials, mainly alkaly compounds of antimony and tellurium. Their QE are in the percentage range, they are sensitive respectively to visible and UV radiation. The drawback is their sensitivity to vacuum conditions.

For the operation of FELs at high charge, high repetition rate and/or long beam pulse the choice of semiconductor photocathodes is mandatory. In the following we present the INFN LASA activity related to development of these kind of photocathodes.

24.2 Cs₂Te Photocathode Production

In our laboratory, after a first experience on alkaly antimonied, we moved to alkaly telluride and in particular on cesium telluride (Cs₂Te) because this photocathode showed better performances and it was more robust than a alkaly antimonied films.

The Cs₂Te production process has been studied by a dedicated R&D activity, in collaboration with Università di Modena [3], where we applied common surface science techniques as AES (Auger Electron Spectroscopy) and XPS (X-ray Photoemission Spectroscopy) for understanding the growth mechanism and the sensitivity of the photocathode to gases exposition. This activity was based on sample compatible with these surface science machines instead of standard plugs used in RF gunns.

These studies showed that our deposition procedure, shown in Fig. 24.1, based on keeping the photocathode substrate at 120 °C while depositing 10 nm of Te and then Cs up to the maximum of the emitted photocurrent, guarantees the right stoichiometric ratio between Cs and Te.

Moreover, we also showed the influence of the substrate on the final photoemissive film composition and its confinement on the surface for Mo substrate, assuring a surface film of few tens of nm well separated from the metal behind.

We then transferred this know-how into the production process based on real Mo plugs, usable in RF gunns.

For the operation in the gun, we must guarantee not only the QE but also its spatial uniformity, low dark current, long life time. In particular the last two points are directly related to the cathode preparation and handling before and during operation in UHV condition that hence needs extreme care. For reducing the dark current, the Mo plug front surface where the photocathode is deposited and that it is exposed to high accelerating fields, was polished to an estimated roughness $R_a \approx 10$ nm [11]. Figure 24.2 shows a plug front surface with a film deposited on top of it.

Fig. 24.1 Quantum efficiency versus cesium thickness. The plateaus correspond to different stoichiometric ratio between Cs and Te [3]

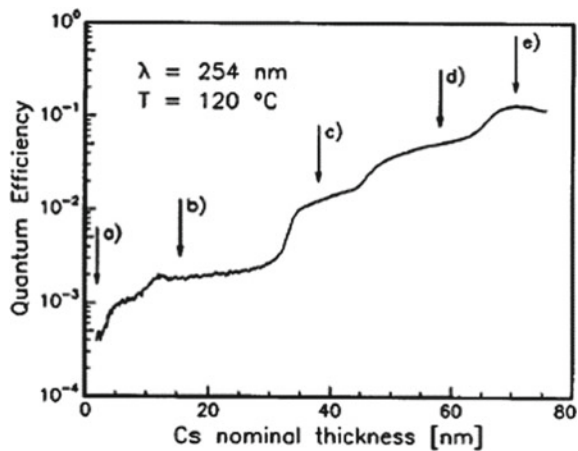
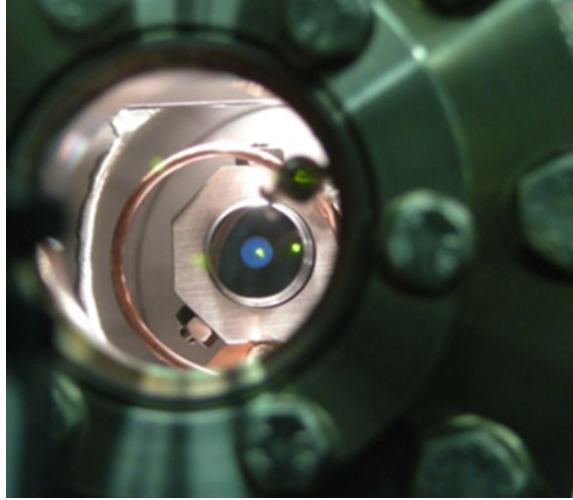


Fig. 24.2 Molybdenum plug with Cs₂Te film deposited on it. The plug is stored in one of our suitcase for maintaining UHV conditions



A critical point has been to assure the proper and reproducible formation of the photoemissive layer on the plugs with the limited diagnostic available in the preparation chamber.

In this system, besides a microbalance for the measurement of the thickness of the evaporated material and a picoammeter for the photocurrent, only optical measurements are possible. For this reason, we have developed dedicated techniques that allows measuring not only photocurrent at the selected wavelength but also the spectral response and reflectivity on-line during the whole production process, spanning a photon energy range from 2.84 to 5.19 eV [10].

The analysis of the photocathode spectral response gives valuable information on its energetic bands and the transition that take place in the material during the photoemission process [6]. A typical QE dependence from photon energies is given by

$$QE = A \left[h\nu - (E_G + E_A) + q_e \sqrt{\frac{q_e \beta F \sin \phi}{4 \pi \epsilon_0}} \right]^m \quad (24.5)$$

where E_G and E_A are respectively the energy gap and the electron affinity of the photocathode, $h\nu$ the photon energy, A is a constant. The third term in the parenthesis represents the contribution from a Schottky effect, relevant mainly for measurements done in RF guns (β field enhancement factor, F electric field at photocathode and ϕ is the emission phase of the electrons w.r.t. the RF phase). Figure 24.3 shows one of such analysis where it is visible a ‘shoulder’ at lower photon energies, usually due to an excess of cesium during the deposition process.

We have also developed a multi-wavelength system that it is nowadays routinely used to monitor the cathode QE during production [10]. This system allows an on-line measurement of the spectral response during the photocathode deposition process.

Fig. 24.3 Spectral response analysis based on Kane model [6]. A shoulder at lower photon energies is clearly visible

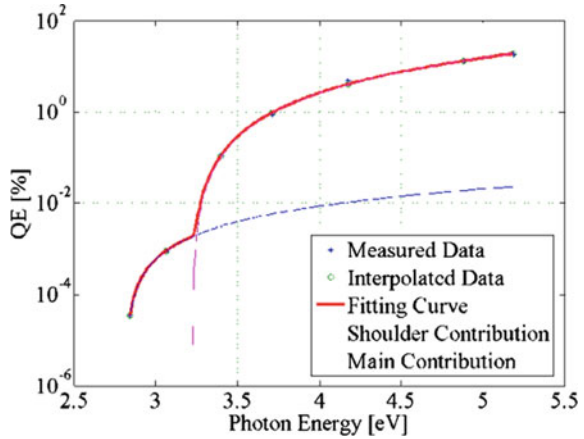
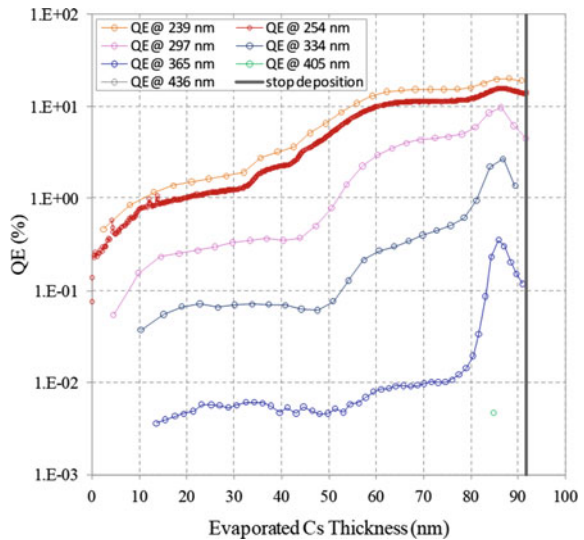


Fig. 24.4 Photocurrent measured at different wavelengths during the cathode formation process. The measured current at longer wavelengths allows detecting the proper formation of the Cs₂Te film



Moreover, this technique allows detecting the proper cathode formation based on a maximum at longer wavelengths as shown in Fig. 24.4.

This assures a better control of the production process and more stable and reliable characteristic of the growth films.

24.3 LASA Photocathodes in the World

INFN LASA is involved in providing photocathodes for operation in RF guns since the '90s. The first system build for this purpose was installed in the A0 experiment at FNAL [2], where the cathode growing chamber was connected directly to the gun



Fig. 24.5 INFN photocathode system connected to the FLASH superconducting accelerator in DESY Hamburg

and the photocathodes where transfer in UHV conditions, to preserve photocathode performances.

Later in the 90s, we developed a split system to physically separate the photocathode production system from the RF gun. The advantage of this solution is the possibility to do photocathode preparation without interfering with the machine operation, very important when the photocathode are used in accelerator dedicated to user as the FELs. The drawback of this solution is the increase complexity of the system that must assure UHV vacuum condition to the photocathodes during all stages of preparation and manipulation from the photocathode deposition to its installation in the gun. The first system of this kind [11] has been installed at the Tesla Test Facility (TTF), now FLASH, at DESY Hamburg (see Fig. 24.5).

Due to the sensitivity of the photocathode to reactive gas exposition, the photocathode have to be kept under UHV condition all the time. A key element is then the transport system and its associated components (carrier, suitcase, etc.). During the years, we have upgraded many of these components to increase the reliability of the system while preserving the photocathode emissive properties. More recently, we have upgraded also the vacuum system of our transport suitcase with the introduction of a new pumping systems based on Non Evaporable Getter (NEG) from a

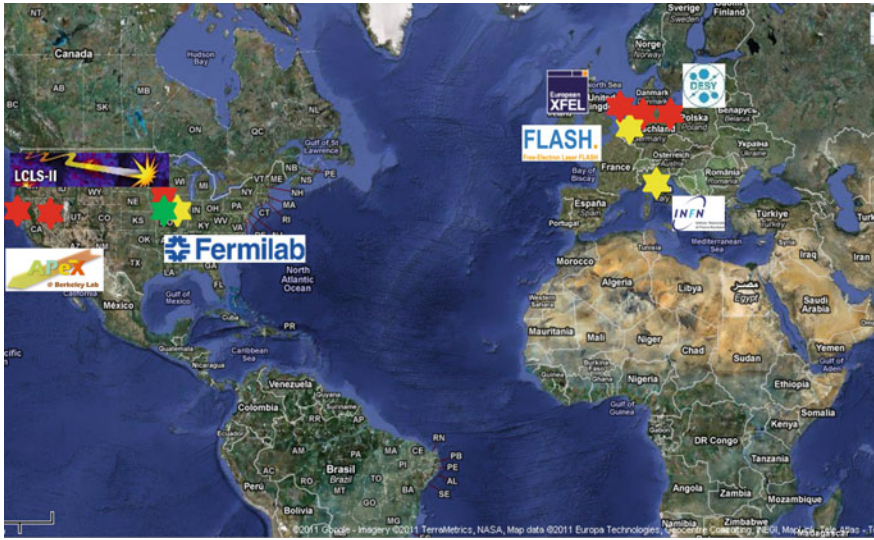


Fig. 24.6 INFN photocathode systems in the world. Green star is a combined system like A0, yellow stars are production systems and red stars are the systems for transferring photocathodes to RF guns

more standard SIP power by a battery car during the long transportation from the production laboratory to the accelerator complex. Besides the advantage of having an active pumping system also without electrical power to guarantee proper vacuum condition, the NEG pump allows also reducing the overall weight of the suitcase, w.r.t. the same configuration with the SIP installed, and shipping the suitcase by plane, not possible with powered SIP [13].

Since the first TTF system, similar systems have been installed in different laboratories around the world: XFEL, PITZ, REGAE, APEX, FAST and for the commissioning and operation of LCLS II (see Fig. 24.6).

The Cs₂Te photocathodes produced up to now are more then 140 with lifetime that now exceed 180 days of continuous operation (24h/24h 7d/7d). The last cathodes have reached an impressive lifetime of more than 400 days of operation, with a peak of more than 900 days for the cathode presently used at FLASH. Figure 24.7 shows the QE of all the cathodes so far produces since 1998, including also the few R&D cathodes grown with different Te thickness. The mean QE of the cathode produced so far is $9.4 \pm 3.2\%$.

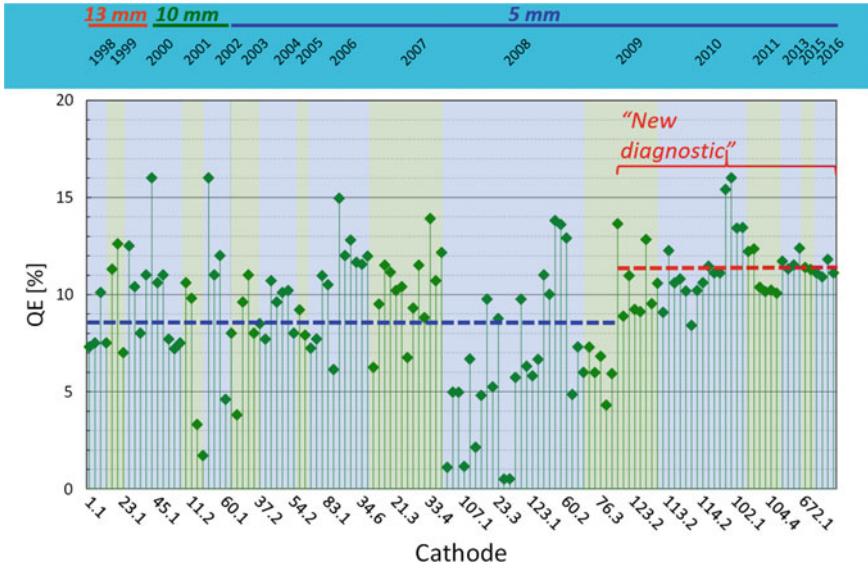


Fig. 24.7 QE of the Cs₂Te photocathodes produced since 1998. On top, we report the diameter of the photoemissive layer reduced from the initial 13 mm to the actual 5 mm. Since the middle of 2009, we have introduced the multi-wavelength techniques that allows improving QE value and reliability

24.4 R&D Activities on Photocathodes

Following the initial R&D activity dedicated to antimony and tellurium alkaly compounds and in parallel with the photocathodes preparation for their use in RF guns, we have continue to investigate the properties of cesium telluride films in view of a deeper understating of the growing mechanism and material characteristics.

24.4.1 Optical Properties

We investigated the optical properties of Cs₂Te by measuring the reflectivity of the film at different angles, polarizations and wavelengths. Being the optical film only few tens of nanometers thick, the contribution of the substrate plays a key role as well as the sample roughness. In Fig. 24.8 a typical measurement of the Cs₂Te film with different polarizations.

The optical parameters have been also used to develop a simple Monte Carlo model [4] of the photoemission process based on the three steps model approach developed by Berglund and Spicer [1]. With this simple model, we have been able to

Fig. 24.8 Angle resolved reflectivity in unpolarized, ‘s’ and ‘p’ polarization for a Cs₂Te film grown with our standard recipe

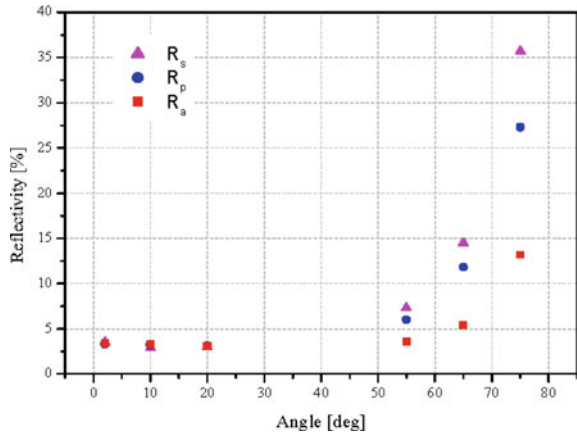
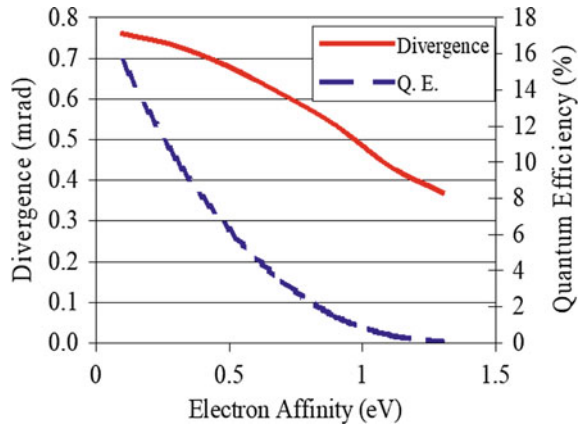


Fig. 24.9 QE and beam divergence ($\propto \epsilon_{\text{therm}}$) simulated using a Monte Carlo model based on our measured Cs₂Te parameters



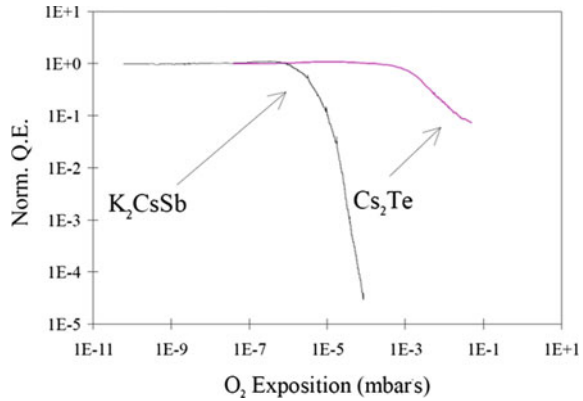
reproduce spectral response curves and to estimate the divergence of the generated electron beam as shown in Fig. 24.9.

24.4.2 Gas Sensitivity and Rejuvenation

As already anticipated, alkaly antimonied and telluride films are sensitive to reactive gas exposition and required UHV condition to preserve their own characteristics. Figure 24.10 shows a comparison between the sensitivity of a K₂CsSb film and a Cs₂Te film to oxygen exposition. As clearly visible, cesium telluride is by at least three order of magnitude less sensitive to oxygen w.r.t. potassium cesium antimonied.

Moreover, Cs₂Te can be rejuvenated, i.e. its QE can be recovered for example after gas pollution by simultaneous illumination with UV light and heating of the

Fig. 24.10 A typical QE degradation due to exposition of a multi alkali antimonied and of a alkaly telluride to O₂ atmosphere. Remarkably, Cs₂Te is at least three order of magnitude less sensitive to this kind of pollution



plug. It is worthwhile to note that both processes (illumination and heating) have to be applied in order for the rejuvenation process to recover the photomissive film QE.

24.4.3 Thermal Emittance

As stated in the introduction, thermal emittance plays a special role in attaining short FEL wavelengths. For this reason we have developed a dedicated Time of Flight apparatus to measure the energy spectrum, angle resolved, of the electrons emitted from a photocathodes [8]. The system is based on a μ -metal chamber, a femto second laser and an electron detector system based on MCP (Multi Channel Plates). The peculiarity of this system is that no external fields are applied to collect the electrons. Since the electron energy ranges from 0 to few eVs, we have dedicated special care in shielding the earth magnetic field and in compensating contact potentials between the sample and the measuring apparatus [9]. Figure 24.11 shows a sketch of the TOF functioning principle.

To reconstruct the thermal emittance from the acquired spectra, we start from calculating the electron kinetic energy:

$$E_{\text{Kin}} = \frac{1}{2} m_0 \left(\frac{L_{\text{TOF}}}{t} \right)^2 \quad (24.6)$$

where m_0 is the electron mass, L_{TOF} is the distance between the sample and the MCP detector and t is the time required by the electron to travel from the sample to the detector. This allows to transform the time of flight to electron kinetic energy. The angle resolved energy spectra are then used to calculate the thermal emittance based on the following formula

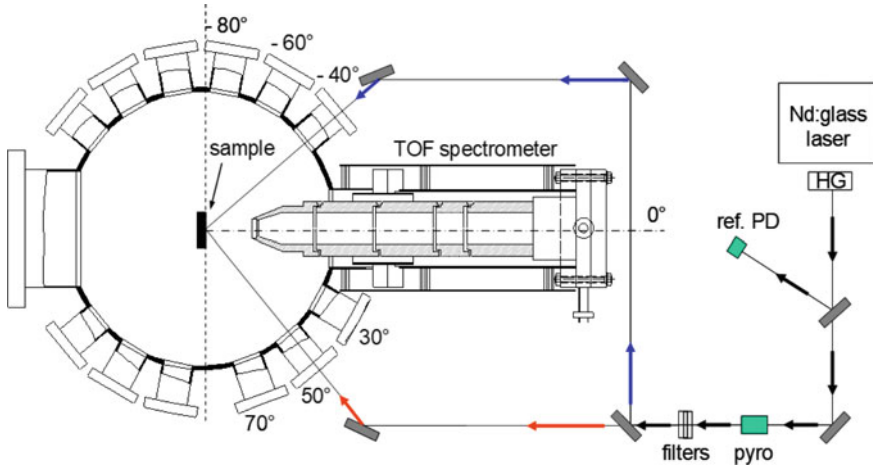


Fig. 24.11 Layout of the time of flight apparatus developed for the measurement of the photocathode thermal emittance

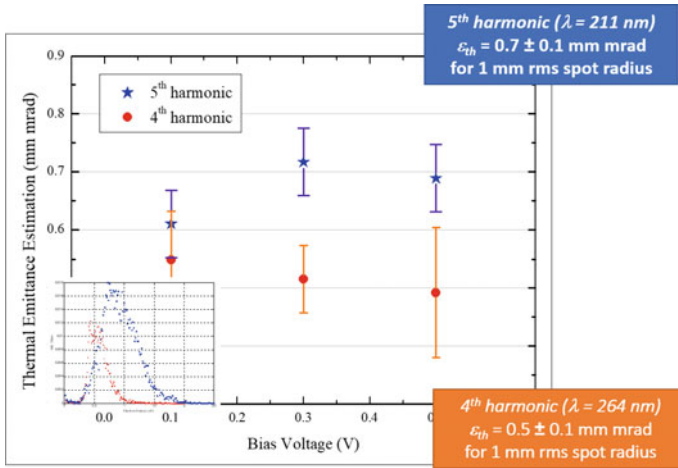


Fig. 24.12 Thermal emittance measurements at the 4th and 5th harmonic of our Nd:Glass laser at different potentials between sample and apparatus for balancing the contact potentials. In the lower left insert an example of electron energy distribution

$$\epsilon_{th} = \frac{1}{2c} \sqrt{\langle r^2 \rangle \left\langle \frac{2 E_{Kin}}{m_0} \cos^2 \theta \right\rangle} \quad (24.7)$$

where it is assumed a spot of 1 mm diameter. Figure 24.12 reports the results obtained for Cs₂Te photocathodes where $\epsilon_{therm}^{4th} = 0.5 \pm 0.1$ mm mrad while $\epsilon_{therm}^{5th} = 0.7 \pm 0.1$ mm mrad [12], corresponding respectively to the 4th and 5th harmonic of our Nd:Glass femto-second laser.

References

1. C.N. Berglund, W.E. Spicer, Photoemission studies of copper and silver: theory. *Phys. Rev.* **136**(4A), A1030–A1044 (1964). <https://doi.org/10.1103/PhysRev.136.A1030>
2. J.-P. Carneiro, et al., Beam transport, acceleration and compression studies in the Fermilab high brightness photoinjector, in *Proceedings of Linear Accelerator Conference* (1998), pp. 878–880
3. A. di Bona, F. Sabary, S. Valeri, P. Michelato, D. Sertore, G. Suberlucq, Auger X-ray photoemission spectroscopy study on Cs₂Te photocathodes. *J. Appl. Phys.* **80**(5), 3024–3030 (1996). <https://doi.org/10.1063/1.363161>
4. G. Ferrini, P. Michelato, F. Parmigiani, A Monte Carlo simulation of low energy photoelectron scattering in Cs₂Te. *Solid State Commun.* **106**, 21–26 (1998). [https://doi.org/10.1016/S0038-1098\(97\)10237-X](https://doi.org/10.1016/S0038-1098(97)10237-X)
5. K. Floettmann, Note on the thermal emittance of electrons emitted by cesium telluride photo cathodes, in *TESLA-FEL Report 1997-01* (1997)
6. E.O. Kane, Theory of photoelectric emission from semiconductors. *Phys. Rev.* **127**(1), 131–141 (1962). <https://doi.org/10.1103/PhysRev.127.131>
7. K.-J. Kim, Brightness, coherence and propagation characteristics of synchrotron radiation. *Nucl. Instrum. Meth. Phys. Res. A* **246**, 71–76 (1986)
8. P. Michelato, L. Monaco, M. Costogliolo, D. Sertore, Thermal emittance estimation using a time-of-flight spectrometer, in *Proceedings of the European Particle Accelerator Conference 2000* (2000), pp. 1669–1671
9. P. Michelato, L. Monaco, D. Sertore, I. Pollini, Time- and angle-resolved photoelectron spectrometer based on a picosecond Nd:glass laser system. *Rev. Sci. Instrum.* **74**(6), 2977–2982 (2003). <https://doi.org/10.1063/1.1574393>
10. L. Monaco, P. Michelato, D. Sertore, C. Pagani, M. Songini, Multiwavelength optical diagnostic during Cs₂Te photocathodes deposition, in *Proceedings of the International Particle Accelerator Conference 2010* (2010), pp. 1719–1721
11. D. Sertore, S. Schreiber, K. Floettmann, F. Stephan, K. Zapfe, P. Michelato, First operation of cesium telluride photocathodes in the TTF injector RF gun. *Nucl. Instrum. Meth.* **A445**, 422–426 (2000). [https://doi.org/10.1016/S0168-9002\(00\)00095-4](https://doi.org/10.1016/S0168-9002(00)00095-4)
12. D. Sertore, D. Favia, P. Michelato, L. Monaco, P. Pierini, Cesium telluride and metals photoelectron thermal emittance measurements using a time-of-flight spectrometer, in *Proceedings of the European Particle Accelerator Conference 2004* (2004), pp. 408–410. [https://doi.org/10.1016/S0168-9002\(00\)00095-4](https://doi.org/10.1016/S0168-9002(00)00095-4)
13. D. Sertore, P. Michelato, L. Monaco, P. Manini, F. Siviero, Use of non evaporable getter pumps to ensure long term performances of high quantum efficiency photocathodes. *J. Vac. Sci. Technol. A* **32**(3), 031602 (2014). <https://doi.org/10.1116/1.4867488>

Chapter 25

LASA Superconducting RF Cavities for Particle Accelerators



Carlo Pagani, Andrea Bellandi, Michele Bertucci, Andrea Bignami,
Angelo Bosotti, Jin Fang Chen, Paolo Michelato, Laura Monaco,
Rocco Paparella, Daniele Sertore, Cecilia Maiano, Paolo Pierini
and Saeid Pirani

Abstract SRF cavities are complex electro-magnetic resonators operating at cryogenic temperature and achieving world-class quality factors and accelerating fields. LASA designed, fabricated and tested many cavities for several research projects.

25.1 Superconducting Cavity Basics

25.1.1 RF Superconductivity

For high duty-cycle operation of particle accelerators, it is mandatory to use Superconducting (SC) cavity to reduce the overall power consumption. The advantage of the SC solution comes from the very low surface resistance of superconductors compared, for example, to standard materials used in RF at room temperature as copper. The widely used superconducting material for cavity fabrication is niobium that has a transition temperature to the superconducting state at $T_c = 9.2$ K.

The dissipated power is related to surface resistance by the following expression

$$P_{\text{diss}} = \frac{1}{2} R_s \int_S |H|^2 dA$$

where H is the magnetic field and S is the cavity surface. In case of normal conducting material, the surface resistance is given by:

C. Pagani (✉)

Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, 20133 Milan, Italy
e-mail: carlo.pagani@mi.infn.it

C. Pagani · A. Bellandi · M. Bertucci · A. Bignami · A. Bosotti · J. F. Chen · P. Michelato
L. Monaco · R. Paparella · D. Sertore · P. Pierini · S. Pirani
INFN - LASA, Via fratelli Cervi 201, 20090 Segrate (Milano), Italy

C. Maiano · P. Pierini · S. Pirani
ESS Accelerator Division, Tunavägen 24, 221 00 Lund, Sweden

© Springer Nature Switzerland AG 2018
P. F. Bortignon et al. (eds.), *Toward a Science Campus in Milan*,
https://doi.org/10.1007/978-3-030-01629-6_25

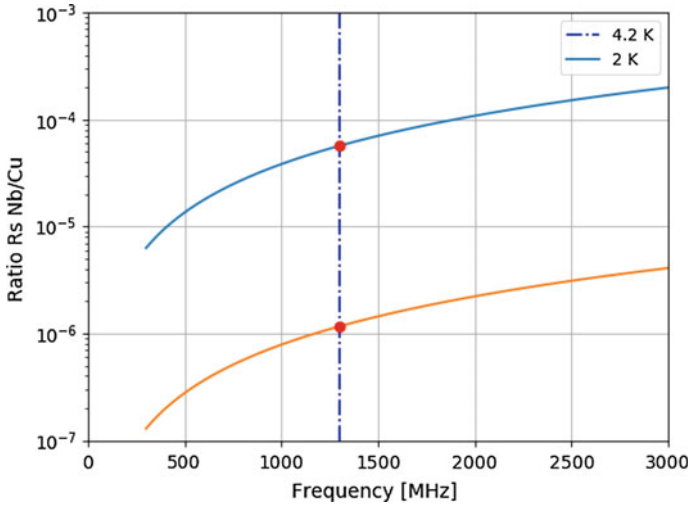


Fig. 25.1 Ratio of Nb and Cu surface resistance versus frequency at different temperatures

$$R_s^{NC} = \sqrt{\frac{\pi \mu_0 f}{\sigma}} = \frac{1}{\sigma \delta}$$

where f is the RF field frequency and σ is the material electrical conductivity. The quantity δ is the skin depth and scales as inverse of the square root of f . For the SC case instead, the surface resistance has a more complex dependence from frequency and temperature and it is drastically reduced by the condensation of conduction electrons into Cooper pairs. A widely used expression for the surface resistance of niobium (Nb) is:

$$R_{BCS} = A \left(\frac{f}{1.5} \right)^2 \frac{1}{T} e^{-\frac{17.67}{T}}$$

where A is a constant dependent upon material characteristics (mean free path, coherence length, London’s depth) and equal to $2 \cdot 10^{-4}$, f is in GHz and T is the material temperature. Figure 25.1 compares the ratio between niobium and copper surface resistance at different temperature for different frequencies showing the advantage of SC at low frequency.

In real cavities one has to consider a residual contribution to the internal surface resistance that is independent from temperature, which accounts for the effect of trapped magnetic field, surface contaminants etc. The total surface resistance can then be written as:

$$R_s^{SC}(T) = R_{BCS}(T) + R_{Mag} + R_{Res}$$

Typical value for $R_{\text{Mag}} + R_{\text{Res}}$ at 1.3 GHz is about $5 \text{ n}\Omega$ (10 mGauss, 2.0 K).

25.1.2 SRF Cavity

The RF cavity provides the electric field needed for accelerating the electron beam. Some figures of merit are useful to describe and compare cavities. The cavity accelerating voltage is defined as the voltage seen by a particle of charge q travelling at the speed of light c on the cavity axis, such that

$$V_{\text{acc}} = \left| \frac{1}{q} \times \text{maximum energy gain possible during transit} \right|$$

The time it takes the charge to cross the accelerating gap must equal half a RF period for the particle to receive the maximum kick from the cavity.

This voltage is given by the line integral of the electric field along the longitudinal cavity axis, E_z , as seen by the particle:

$$V_c = \left| \int_0^{L_{\text{cav}}} E_z(\rho = 0, z) e^{i\omega_0 z/c} dz \right|$$

where L_{cav} is the cavity active length and $\omega = 2\pi f$ is the angular frequency of the accelerating mode. Often one quotes the average accelerating electric field E_{acc} , that the particle sees during its transit. This is given by:

$$E_{\text{acc}} = \frac{V_{\text{acc}}}{L_{\text{cav}}}$$

The first fundamental parameter is the quality factor of the cavity resonator that relates the stored energy to the dissipated power per RF cycle:

$$Q_0 = \frac{\omega_0 U}{P_{\text{diss}}}$$

where U is the electromagnetic stored energy and P_{diss} is the dissipated power. To relate the dissipated power P_{diss} to the accelerating voltage V_c , it is useful to introduce the shunt impedance R such that:

$$P_{\text{diss}} = \frac{V_{\text{acc}}^2}{R}$$

At this point, we introduce two figures of merit which are dependent only from the cavity geometrical parameter and not from its RF properties. The first is the ratio between R and Q_0 , defined as:

$$\frac{R}{Q_0} = \frac{V_{\text{acc}}^2}{\omega_0 U}$$

Since both ω_0 and the ratio V_{acc}^2/U scale inversely with the cavity linear dimensions, the parameter “R over Q” is independent of the cavity size. The quantity “R over Q” is also often used for determining the level of the mode excitation by charges passing through the cavity. Now it is easy to see that the power dissipated by the cavity at a given accelerating gradient is:

$$P_{\text{diss}} = \frac{(E_{\text{acc}} L_{\text{cav}})^2}{\frac{R}{Q_0} Q_0}$$

The second useful figure of merit is the “geometrical factor” G which relates the quality factor Q_0 and surface resistance R_s by the following expression

$$Q_0 = \frac{G}{R_s}$$

Once the cavity is equipped with its antennas (main coupler antenna, pick up antenna, etc.) new dissipations are added to the system. The coupling factor corresponding to each port x is then given by:

$$\beta_x = \frac{P_x}{P_{\text{diss}}} = \frac{Q_0}{Q_x}$$

where P_x is the power dissipated by the x component. In a realistic case where only main coupler (MC) and pick up (PU) antenna are installed, the “loaded” quality factor Q_L is related to the cavity quality factor Q_0 :

$$Q_0 = Q_L(1 + \beta_{\text{MC}} + \beta_{\text{PU}})$$

25.2 Electron Cavities, TESLA Technology

Resulting from a global effort of the SRF community and DESY in particular since the late 90s, the TESLA-type cavity [1] (Fig. 25.2, Table 25.1) has been the core element of many projects so far: TTF, FLASH and then European-XFEL at DESY, ILC, LCLS-II at SLAC, SLS at Shanghai, just mentioning the largest ones.

The LASA team from INFN Milano strongly participated to the activity that led to its final design, its fabrication as well as the production of the numerous tools related to its manufacturing, working in a strong connection with industrial partners.

Recently, for the European-XFEL project based in Hamburg (Germany) LASA has been in charge for the production of 50% of the 800 T resonators of the E-XFEL linac.



Fig. 25.2 The TESLA-type superconducting cavity

Table 25.1 Key parameters for the TESLA-type cavity

Parameter	Value
Type	Standing wave, bulk niobium
Accelerating mode	TM010, π mode, 1300 MHz
Number of cells	9
Total length	1247 mm
Effective length	1038 mm
Iris diameter	70 mm
Cell-to-cell coupling	1.87%
R/Q	1036 Ω
Geometry constant	270 Ω
$B_{\text{peak}}/E_{\text{acc}}$	4.26 mT/MV/m

A fundamental inheritance of the joint effort of LASA and DESY teams was indeed that the vast knowledge gathered at different labs has been transferred to few, highly specialized, European industries. As a result these companies were, eventually, capable of delivering the full production cycle, from raw material to the dressed cavity ready for cold test.

The E-XFEL series has been the first large-scale production of TESLA cavities and nonetheless the demanding specifications of the project were successfully met with margin. To be accepted for the installation in the linac indeed an “usable” accelerating gradient $E_{\text{usable}} \geq 23.6$ MV/m and an intrinsic quality factor $Q_0 \geq 1 \cdot 10^{10}$ were required, where “usable” definition includes additional operational parameters as x-ray field emission and Q_0 performances [8].

A statistical analysis of the large amount of data collected during the experimental qualification campaign of the E-XFEL cavity series [2] is presented in the two following plots (Figs. 25.3 and 25.4), confirming how this activity, led by DESY and LASA, defined the state-of-the-art for this superconducting technology.

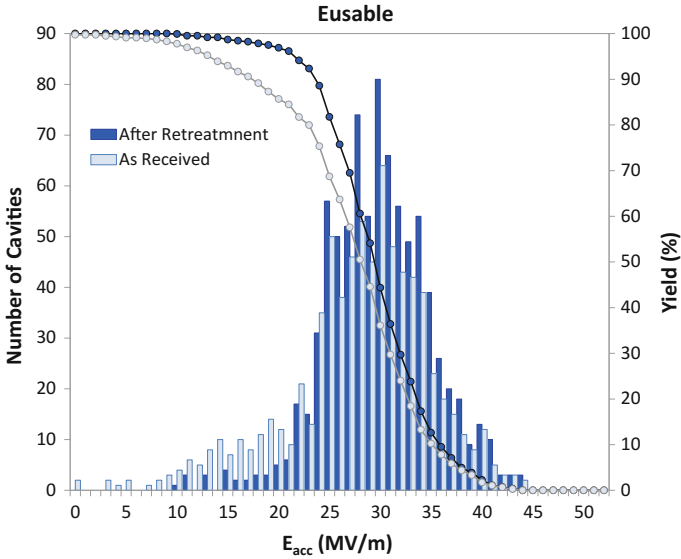


Fig. 25.3 Yield versus Performance analysis for the entire E-XFEL production. The average usable gradient, after retreatments, resulted to be 29.8 ± 5.1 MV/m

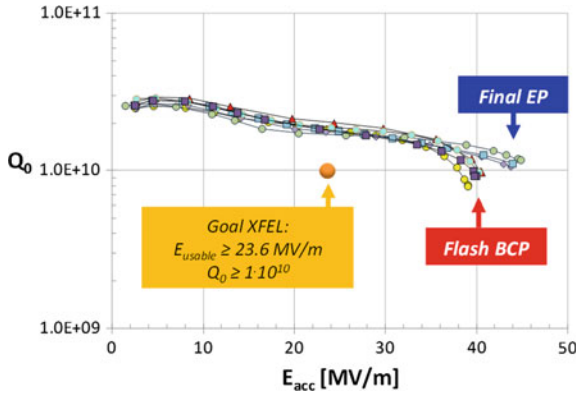


Fig. 25.4 Q_0 versus E_{acc} plot for the subset of the best performing resonators. The different final surface preparation procedure held in the two producing companies was indeed impacting the peak field value (Final Electro-Polishing at RI, Germany, and Flash Buffered Chemical Polishing at E. Zanon, Italy)

Alongside cavities, LASA contributed through the years to the development of many different aspects of the SRF technology. During the early TESLA activities, LASA had a leading role in the design and fabrication with industry of the large and modular cryostats, often called cryomodules, that host the cavities in the linac in operating conditions.

Furthermore and among many others, a wire-position monitor (WPM) [3] for the analysis of the cryomodule cold-mass strain during thermal cycles and a novel cold tuning system coaxial to the cavity (Blade Tuner) [4] were both developed by LASA team and successfully operated.

25.2.1 3.9 GHz European-XFEL Injector Cavities

In addition to its leading role in the cavity series production, LASA team contributed to E-XFEL by designing and realizing the third-harmonic cryomodule for the linac injector. This module, one of a kind in the SRF projects scenario, hosts eight 9-cells cavities operating at the third harmonic of the linac RF systems (thus 3.9 GHz).

Most of the key components of the module system were a LASA contribution:

- The design of these small cavities, their prototyping, fabrication and experimental qualification.
- The helium vessel and its magnetic shielding, designed to be compatible to the coaxial tuning action.
- Cold cavity tuning systems, coaxial to the resonator and inspired by the INFN Blade tuner already developed for the TESLA-type cavity.
- The design and realization of the whole cryogenic system, vessel, shielding, piping and vacuum fittings.

Specifically concerning the cavities, LASA directly provided all the steps throughout the life-cycle of these novel and unique items:

- Design, prototyping and RF measurements at room temperature.
- Definition and responsibility of the surface treatments to be done at the industrial partner premises.
- Cleaning and preparation of each cavity at the LASA class-10 clean-room, including the high-pressure surface rinsing with ultra-pure water (HPR), the installation of RF ancillaries and final slow-pumping.
- Qualification of each cavity by means of a cryogenic test (2.0 K super-fluid helium bath temperature) in the LASA vertical cryostat.
- Assembly of cavity string in the cryomodule at DESY.

All the 20 third-harmonic cavities (Fig. 25.5) eventually produced (including those for a spare injector module identical to the first one) outperformed the specifications required [5]: $E_{acc} \geq 15$ MV/m and intrinsic quality factor $Q_0 \geq 1 \cdot 10^9$ (Fig. 25.6).

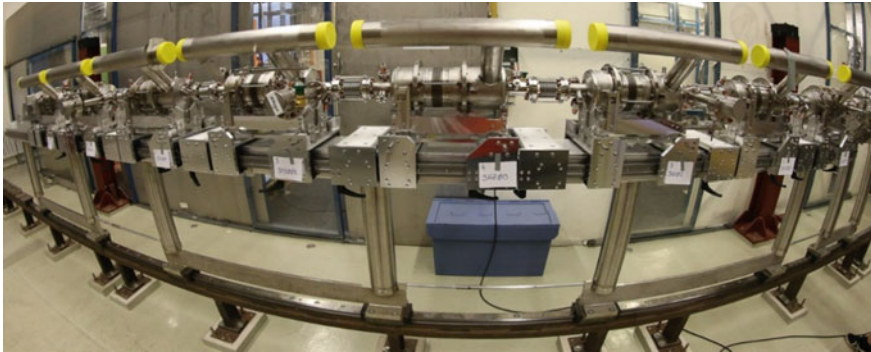


Fig. 25.5 String of 3.9 GHz cavities just after assembly and roll-out from clean-room at DESY, intermediate stage during third-harmonic cryomodule assembly

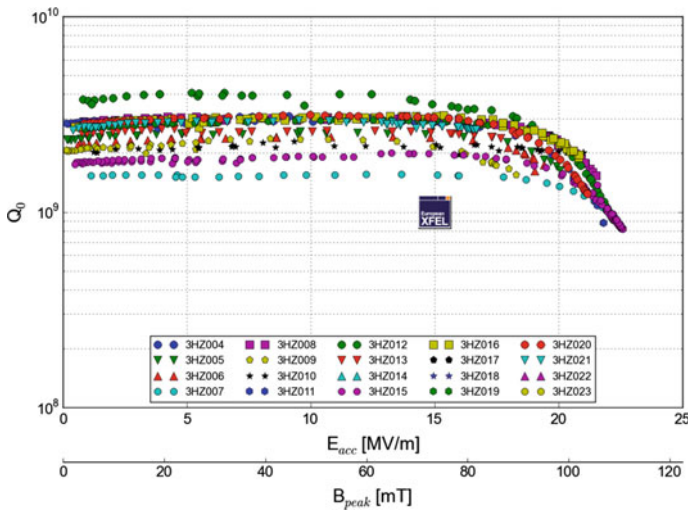


Fig. 25.6 Summary plot for the whole set of 20 3.9 GHz cavities produced, the E-XFEL logo marks the project specifications

25.3 Cavities for Protons

The global success of the joint effort for the TESLA-type cavity, triggered novel applications for the SRF technology applied to protons. Several projects worldwide pursued the benefits of superconducting linacs to key topics as the transmutation of nuclear wastes, accelerator-driven sub-critical reactors (MYRRHA project in Belgium, C-ADS in China) and neutron production by spallation (SNS in US, and ESS in Sweden).

Differently from electron machines, the proton acceleration profile along the linac requires the use of many different accelerating structures, each covering a limited range of beta values (where $\beta = v/c$). Elliptical multi-cell cavities were successfully operated at first in the high-beta regime (from 0.6 to 1), while LASA team pioneered the lower-end region by realizing two, one of a kind in its design, beta 0.47 5-cell cavity specifically designed for the acceleration of high-intensity proton beam [6].

An innovative software for the design of elliptical cavities was conceived at the time at LASA with the specific purpose of defining the geometry of this low-beta cavity. As of today, indeed, this tool (named “BuildCav”) is still a widely used support in several laboratories.

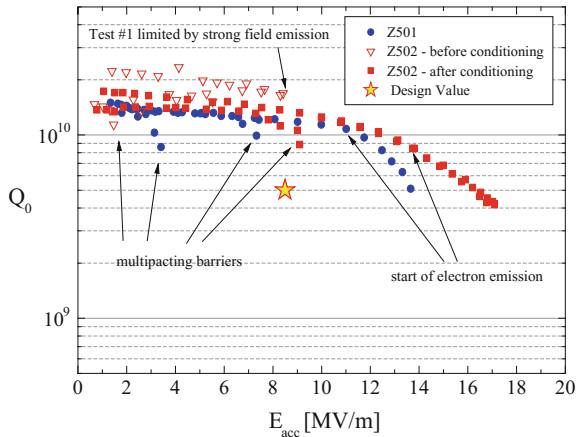
Both prototypes were qualified in operating conditions in 2004 by means of vertical cold tests at different laboratories and both outperformed their nominal specifications. Cavity Z502 in particular proved, after final surface conditioning, remarkable peak surface electric and magnetic field measured as 61 MV/m and 100 mT respectively.

This successful step made these resonators available for the following activity in the framework of many international projects (Figs. 25.7 and 25.8).

Fig. 25.7 LASA beta 0.47 cavity prototype



Fig. 25.8 Vertical cold tests results for the two LASA beta 0.47 cavities, in TJNAF (US) for the Z501 and in CEA Saclay (France) for the Z502



Alongside low-beta resonators, LASA significantly contributed to the feasibility study of a high-intensity proton linac by developing all the ancillaries required by a fully functional cavity (a “dressed” cavity) and more:

- Fully operative and on specs prototypes of cavity ancillaries:
 - Piezo-tuners: coaxial type inspired to the Blade Tuner of TESLA cavities, equipped with two piezo-ceramic actuators to compensate rapidly-varying detuning.
 - Magnetic shielding: innovative cryo-perm shield installed within the helium tank to minimize residual field as well as space occupancy around the cavity.
 - Helium tank, support pads and rollers, etc.
- Beam dynamic, beam halos and machine reliability studies.
- Cryomodule design and realization of a short prototype hosting a single cavity with a tie-rod/space-frame suspension scheme and a bottom-up power coupler.

In the framework of the collaboration between LASA and IPN Orsay (France), this latter single-cavity prototype cryomodule was finally made fully functional in 2013 by installing it, together with the dressed 5-cells Z502 cavity, in the cryo-test area at IPN equipped with cold-box and 80 kW CW IOT RF amplifier.

This prototype modular unit was then successfully commissioned [7] and it still serves as a unique technical test bench for high-intensity proton linac issues, addressing for instance the study of reliability-oriented control system and fast fault recovery procedures (Figs. 25.9 and 25.10).

After this pioneering phase, LASA team is currently going to develop and produce new cavities in the framework of other proton-related projects as the PIP-II at Fermilab (US) and the ESS in Lund (Sweden) described in the following chapter.

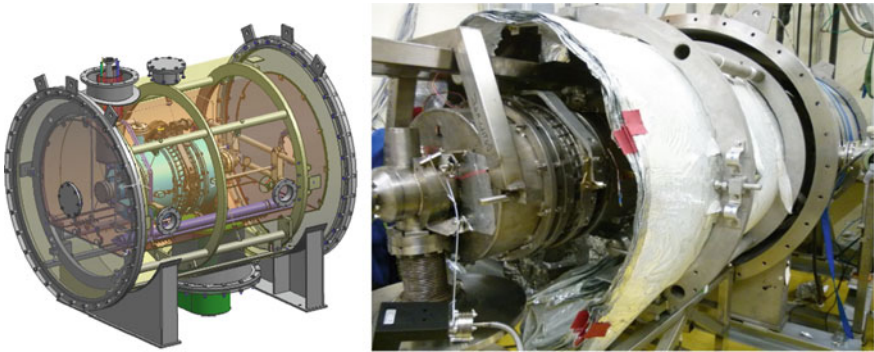


Fig. 25.9 Artistic view of the complete single-cavity module (left) and the installation of Z502 cavity at IPN (right). The Blade piezo-tuner is also visible, coaxially installed in the middle of the helium tank

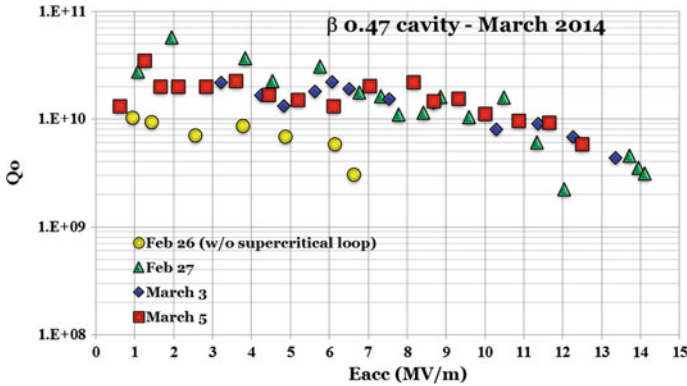


Fig. 25.10 Q versus E_{acc} plot for the Z502 cavity with high-power coupler at the single-cavity cryomodule test stand in IPN Orsay

25.3.1 European Spallation Source—ESS

Since early 2015, the LASA team at INFN Milano is in charge of the in-kind contribution of the whole medium beta section of the proton linac of the ESS, a pan-European neutron source project that will be the largest accelerator facility built in Europe after European-XFEL and will host the most powerful superconducting proton linac ever built.

In details, the LASA contribution extends to:

- The procurement of niobium for the fabrication of cavities.
- The fabrication of the series of 36 cavities (plus 2 spares) including surface treatments, frequency tuning and integration in the helium tank. The full package of manufacturing and treatments will be delivered by the vendor industry.
- Certification activities, documents and ancillaries management.
- Qualification cold test in a qualified infrastructure (DESY, Germany).
- The cavity of the series is going to be delivered at the ESS cryomodule installation site at CEA Saclay by mid-2018, followed by all the others at a rate of four cavities per month.

In view of the final series production, an intense R&D activity is on-going at LASA (Figs. 25.11 and 25.12), two prototypes have been designed (Table 25.2), built and cold-tested and two different materials have been exploited (fine-grain and large-grain niobium). As for the final cavities, the full sequence of surface treatments has been conceived by LASA and provided by the manufacturing industry.

The LASA ESS-MB cavity is currently installed in the first medium-beta demonstrator cryomodule of the ESS linac undergoing experimental qualification at CEA-Saclay.

Table 25.2 Key parameters for the LASA ESS-MB cavity

Parameter	Value
Type	Standing wave, bulk Niobium, $\beta = 0.67$
Accelerating mode	TM010, π mode, 704.42 MHz
Number of cells	6
Effective length	0.855 mm
Iris diameter	50 mm
Optimum beta	0.705
Cell-to-cell coupling	1.55%
R/Q at opt. beta	374 Ω
Geometry constant	200 Ω
B_{peak}/E_{acc}	4.95 mT/MV/m



Fig. 25.11 LASA ESS-MB prototype cavity during chemical etching of the RF internal surface via BCP (Buffered Chemical Polishing). The outer surface temperature has been monitored during this treatment by means of an infrared camera (small frame in the bottom)

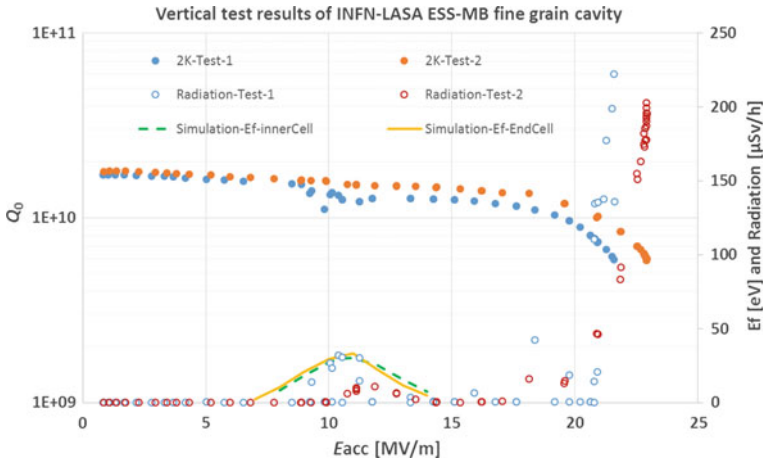


Fig. 25.12 Results of LASA ESS-MB cavity vertical cold-test performed at LASA cryostat. Although this was the first prototype built it outperformed with margins the project requirements of $E_{acc} \geq 16.7$ MV/m and intrinsic quality factor $Q_0 \geq 5 \cdot 10^9$

References

1. B. Aune et al., *The Superconducting TESLA Cavities*. [arXiv:physics/0003011](https://arxiv.org/abs/physics/0003011) (2000)
2. D. Reschke et al., Performance in the vertical test of the 832 nine-cell 1.3 GHz cavities for the European X-ray free electron laser. *Phys. Rev. Accelerat. Beams* **20**, 042004 (2017)
3. A. Bosotti et al., On line monitoring of the TTF cryostats cold mass with wire position monitors. In INFN/TC-00/02 (2000)
4. R. Paparella, Ph.D. Thesis (2008). cds.cern.ch/record/1118577/files/care-thesis-08-001
5. P. Pierini et al., Fabrication and vertical test experience of the European X-ray free electron laser 3.9 GHz superconducting cavities. *Phys. Rev. Accelerat. Beams* **20**, 042006 (2017)
6. P. Pierini et al., Low beta elliptical cavities for pulsed and CW operation. In *Proceedings of PAC09* (Vancouver, BC, Canada, WE5PFP034)
7. F. Bouly, Ph.D. Thesis (2011). <http://tel.archivesouvertes.fr/tel-00660392>
8. W. Singer et al., Production of superconducting 1.3-GHz cavities for the European X-ray free electron laser. *Phys. Rev. Accelerat. Beams* **19**, 092001 (2016)